

69257

192545

m9

STIC-Biotech/ChemLib

From: Vogel, Nancy  
Sent: Monday, June 12, 2006 9:13 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search 09/896,888

please do a sequence search, including interference files, of

SEQ ID NO:1 (entire),  
bp. 351-527 of SEQ ID NO: 1,

of 09/896888, and return results to me on paper asap.

Thanks  
Nancy Vogel  
Patent Examiner  
Art Unit 1636  
Office: Remson 2A70  
Mail Box: Remson 2C70  
(571) 272-0780

Nancy Vogel  
Patent Examiner  
Art Unit 1636  
Office: Remson 3C09  
Mail Box: Remson 2C70  
(571) 272-0780

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIS: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



Db	1	CATGATGATAAACAAATGATGATGCTAATGTTGCTTCAACAAACAATCTGTTGAACCTG	60
Qy	61	TTTTCAATGTTTCCCAACAAGCAGCTTTATACCTGGTGGCTTCCCAACAACAATCTTTT	120
Db	61	TTTTCAATGTTTCCCAACAAGCAGCTTTATACCTGGTGGCTTCCCAACAACAATCTTTT	120
Qy	121	GCACATGCAAAAAACAGCTTTTGGACGGGGCCCATCATAGTACAAACTTACGTTTC	180
Db	121	GCACATGCAAAAAACAGCTTTTGGACGGGGCCCATCATAGTACAAACTTACGTTTC	180
Qy	181	GTAAGACTATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACAC	240
Db	181	GTAAGACTATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACAC	240
Qy	241	ATTGAACCTTTTGGACAGTGCAAAAAAGTACGTTGGAGAGTCAACGTAGCGCGGCTTATC	300
Db	241	ATTGAACCTTTTGGACAGTGCAAAAAAGTACGTTGGAGAGTCAACGTAGCGCGGCTTATC	300
Qy	301	GGGTGCGGCTCTGTCAAGTACGAATCACTTATGGAACCGACAGAGTGTGTTATATGT	360
Db	301	GGGTGCGGCTCTGTCAAGTACGAATCACTTATGGAACCGACAGAGTGTGTTATATGT	360
Qy	361	GACAGGAACGCAGGCTCCGTGTGTCCTTAAACGACCGACCGGACCACTCCTTATGGGAACA	420
Db	361	GACAGGAACGCAGGCTCCGTGTGTCCTTAAACGACCGACCGGACCACTCCTTATGGGAACA	420
Qy	421	GGACGCGCTCATATACAGCGCGGCTTATCTCATGCGCGTGAACCGGACAAGAGGCGCC	480
Db	421	GGACGCGCTCATATACAGCGCGGCTTATCTCATGCGCGTGAACCGGACAAGAGGCGCC	480
Qy	481	GTCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGTGTAACAAGTTGAACAG	540
Db	481	GTCGCGCTTATCGCGCTTATAAATACAGCCCGCAACGATCTGTGTAACAAGTTGAACAG	540
Qy	541	CATCTGTTACAGCGCAACAACATG 564	
Db	541	CATCTGTTACAGCGCAACAACATG 564	

RESULT 2	
NPHTTAA	
LOCUS	4170 bp DNA linear VRL_02-AUG-1992
DEFINITION	Orygia pseudotsugata nuclear polyhedrosis virus transcriptional trans-activator (IE-2) gene, complete cds; ORF, complete cds.
ACCESSION	M83827.1 GI:332540
VERSION	transcriptional transactivator.
KEYWORDS	Orygia pseudotsugata single capsid nucleopolyhedrovirus Orygia pseudotsugata single capsid nucleopolyhedrovirus Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Nucleopolyhedrovirus.
SOURCE	1 (bases 1 to 4170)
ORGANISM	Theilmann,D.A. and Stewart,S.
REFERENCE	Molecular analysis of the trans-activating IE-2 gene of Orygia pseudotsugata multicapsid nuclear polyhedrosis virus
AUTHORS	Virology 187 (1), 84-96 (1992)
TITLE	1736546
JOURNAL	2 (bases 1 to 4170)
PUBMED	Theilmann,D.A. and Stewart,S.
REFERENCE	Tandemly repeated sequence at the 3' end of the IE-2 gene of the baculovirus Orygia pseudotsugata multicapsid nuclear polyhedrosis virus is an enhancer element
AUTHORS	Virology 187 (1), 97-106 (1992)
TITLE	1736547
JOURNAL	Original source text: Orygia pseudotsugata nuclear polyhedrosis virus DNA.
PUBMED	
COMMENT	

FEATURES	source	location/Qualifiers
gene		1..4170 /organism="Ornyia pseudotsugata single capsid nucleopolyhedrovirus" /mol_type="genomic DNA" /db_xref="taxon:10450" 627..3959

[illegible]

ORIGIN	Query Match	100.0%;	Score 564;	DB 10;	Length 4170;
	Best Local Similarity	100.0%;	Posed. No. 2,4e-185;		
	Matches 564;	Conservative 0;	Mismatches 0;	Gaps 0;	
QY	1	CATGATGATTAACCATGATGATGCTGCTGCTTCAACACAACTTCGTGTAACGTG	60		
Db	130	CATGATGATTAACCATGATGATGCTGCTGCTTCAACACAACTTCGTGTAACGTG	189		
QY	61	TTTTCAGTTGGCCAAAGACCTTATATCTGGTGGGCCCTCCACACCAACTTTTT	120		
Db	190	TTTTCAGTTGGCCAAAGACCTTATATCTGGTGGGCCCTCCACACCAACTTTTT	249		
QY	121	GCACTGCAAAAAAACAACGCTTTTGGCAGCGGGCCCATATAGTACAACTTACGTTTC	180		
Db	250	GCACTGCAAAAAAACAACGCTTTTGGCAGCGGGCCCATATAGTACAACTTACGTTTC	309		
QY	181	GTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCCAAATPACCTACGACAC	240		
Db	310	GTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCCAAATPACCTACGACAC	369		
QY	241	ATTGACCTTTTGGCAGTGCACAAAAAGTACGTCGTGGCAGTCACTGATGCGCGCTTATC	300		
Db	370	ATTGACCTTTTGGCAGTGCACAAAAAGTACGTCGTGGCAGTCACTGATGCGCGCTTATC	429		
QY	301	GGGTGCGGTCCTGTCAAGTACGAATTCATTTATGGACCGGACGAGTGTGTTATCGT	360		
Db	430	GGGTGCGGTCCTGTCAAGTACGAATTCATTTATGGACCGGACGAGTGTGTTATCGT	489		
QY	361	GACAGAACGCAAGCTTCTGTGTTGCTTACCGCAGCGGACGCACTTCCTTATGGAAACA	420		





/db\_xref="GI:1911255"  
/translation="MYEASOIDEHFVGYGDNEMALRETFENHATASVISLSDPA  
PIRHATGLPGDHIHYCEAPPCALPNMPLXYDMVPIGREGKRVYHVCASGR  
SALAVYTLKRSROMPYQDALNAVOSKRRVALINDHFRPLATRCSTRFVNDLKPQIV  
"

CDS

complement(6129..6791)

/note="ORF10; ptp-1; similar to AcMNPV ORF1"

/codon\_start=1

/product="protein tyrosine phosphatase 1"

/protein\_id="AAC5909.1"

/db\_xref="GI:1911256"

/translation="MLTSRRMAVITYRTSRNCGAWCTINGVRRRQVHLPSVRAPM  
FPRMHEYTAGVIEGTRLLCFKVPILNAELFEVYSDERMTAATLASHSLGAVT  
DLNTARYDGAQVQKMLYKRIYGVRAIPDDIYAELETVDDEFRCSPMLVAV  
HMHGLNRSGLVLCRVVVERLVSPDIALRFEPAHGHKIERITNYIQDLLARHNVGO  
PN"

CDS

6819..7814

/note="ORF11; similar to AcMNPV ORF11"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59010.1"

/db\_xref="GI:1911257"

/translation="WLSLSKLVYAYGYGNLPHDRYGESYHLRYIVHEHLNTVYSN  
ASCVRDIATARCINSGHLCFVARQLDVS EVAARLSAARFGDAGLCAHQRLA  
DIRHAPLARVGRANIPALDIADIPSDVTNNLQGIIGRFMHFPGSGIARVADVF  
DPDIRADGMWYHKECVLTVMHLYVACGAPAGSATRLDVAKAIIPNDEGNCAPALAA  
VYGRFCAIGREHFAHHTKACMHILFQMRNDLPADRHPCFGVITKDFGRCKDTYD  
LRTHADALYHGTTRDROKMLFDLCCVNASDIDADQYDCVANKFYATQNKYK"

CDS

complement(8180..9142)

/note="ORF12; similar to AcMNPV ORF13"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59011.1"

/db\_xref="GI:1911258"

/translation="MFSMFGMWNADDEQVNAERDEQAYRYIVDQRAHSDLYRDFV  
RCHPVPKRYVLDETAGRCSSVDFCKGLKISHDLQRCNFRQHRQNLBELVLAAP  
PAQPDLSGLFATGHGLVOLLOQFSPANKKEVLLAVGANDHNDMLDGIKEVLNAP  
KTLNTNDSKFISSAKSFKEVGAARFQEDRLDTRKLNALNCCAPATRPAPVPER  
DVTNPHLAIFMGVEDRGVLTQIAFARGSEHFRKRLREBEGNDVVRARNPILA  
VHCKEERFANGKHIRLPRKVIIEVDCAVNAVADIVKALINT"

CDS

complement(9094..9825)

/note="ORF13; left-1; similar to AcMNPV ORF14"

/codon\_start=1

/product="late expression factor 1"

/protein\_id="AAC59012.1"

/db\_xref="GI:1911259"

/translation="MAFCKYTPERVQMMALAYNDSRLAFTDPRRYVHANNFDS  
AAQIFATVKNISIDVAVKPLEEGGRWVIDADPKDCAFAELMLKVNAGATAFMLF  
EGGEDAVQRTMFSGNRGFLMLKFCGFKNDAPKSLREHNPVQKPAKLVSGDIRPG  
SFADCVRAVHMYIGDAREDLVLRYPDVDRVCANMKGIRAPFSYNGKDYSRCL  
TOOLOORIKASAGCLAGTPTPSK"

CDS

9893..11362

/note="ORF14; egt; similar to AcMNPV ORF15"

/codon\_start=1

/product="ecdysteroid UDP-glucosyltransferase"

/protein\_id="AAC59013.1"

/db\_xref="GI:1911260"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

CDS

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

CDS

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

CDS

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

CDS

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

CDS

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

/product="unknown"

/protein\_id="AAC59014.1"

/db\_xref="GI:1911261"

/translation="MVFLLITLTLTLAGARASLAVLPTPAYSHVYVYAVYALVK  
NCHNVYIKQQLLDYAVQDECGREVEDADMSAQYKUYASSGVPRKRVAVADETTV  
TANVYGLIEMFTQDFDNANVRPLFNTRTFDAVVVEAFDVALVEGHLRPAVIOI  
APGYGLIEMFERRAVAPRLPHYPTEGAALTRGGALISMRLLNEBELARSDEL  
KQOFGKSTPTIROLRQNVQQLLNLHVPYNNRPVPSVQYGGGLHQAQLPQRLDA  
PLERRLNSVDGAVYVFGSGIDTNSIHAEFLMLDPTPANLNNTYTLAKVADVAAS  
VALPRNYLAQKFSQTAIVLRHKVAVAFVAGOSSEALQRAVAVPVCILPMMDOPFH  
ARLQCGVAVARALDTAVASAPQALAREYIADGEARAIIDLRAVEHDAAPDEKA  
VKTERVYIKFNNDVMPARSLKTYANMAASIDVFVRPL"

11547..12140

/note="ORF15; similar to AcMNPV ORF16"

/codon\_start=1

## CDS

763..990  
 /gene="p8.9"  
 /note="8.9 kda basic protein"  
 /codon\_start=1  
 /protein\_id="AAB27738.1"  
 /db\_xref="GI:404519"  
 /translation="MNSMOKIRMAKQOVVARQHRRAKLGRVYKAKLRAELCEKIQ  
 LQRVNDALAKAFEEFVYPHFSFYLTLN"

## ORIGIN

Query Match 99.7%; Score 562.4; DB 10; Length 1429;  
 Best Local Similarity 99.8%; Pred. No. 8.2e-185;  
 Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATGATGCTAATGCTTCAACAACAATCTGTTGAACGTG 60  
 DB 765 CATGATGATTAACAATGATGCTAATGCTTCAACAACAATCTGTTGAACGTG 706  
 QY 61 TTTTCATGTTTGCACACACCTTTTATCTCGTGCGCTCCCAACCAACTTTT 120  
 DB 705 TTTTCATGTTTGCACACACCTTTTATCTCGTGCGCTCCCAACCAACTTTT 646  
 QY 121 GCACGTGCAAAAAACAGCTTTTGCACGCGGCCCATATAGTACAACTTACGTTTC 180  
 DB 645 GCACGTGCAAAAAACAGCTTTTGCACGCGGCCCATATAGTACAACTTACGTTTC 586  
 QY 181 GTAGACATTTTATCATTAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 240  
 DB 585 GTAGACATTTTATCATTAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 526  
 QY 241 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTGCGGAGTCAAGTACGCGGCTTATC 300  
 DB 525 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTGCGGAGTCAAGTACGCGGCTTATC 466  
 QY 301 GGGTCGCGTCTGTACGTAAGTACATTAATGACCGGACGAGTGTCTTATCGT 360  
 DB 465 GGGTCGCGTCTGTACGTAAGTACATTAATGACCGGACGAGTGTCTTATCGT 406  
 QY 361 GACAGACGCGCAGCTTCGTGTGTTGCTAACCGGACCGGACGCACTCCCTTATCGAACA 420  
 DB 405 GACAGACGCGCAGCTTCGTGTGTTGCTAACCGGACCGGACGCACTCCCTTATCGAACA 346  
 QY 421 GGAAGCGCTTCATATCAGCGCGGCTTATCTCATGCGCGTGAACGGAACGAGCGCC 480  
 DB 345 GGAAGCGCTTCATATCAGCGCGGCTTATCTCATGCGCGTGAACGGAACGAGCGCC 286  
 QY 481 GTCCCGCTTATCGCGCTATTAATATACGCGGACGATGTGTAACACAGTTGAACAG 540  
 DB 285 GTCCCGCTTATCGCGCTATTAATATACGCGGACGATGTGTAACACAGTTGAACAG 226  
 QY 541 CATCTGTTACAGGACCAACATG 564  
 DB 225 CATCTGTTACAGGACCAACATG 202

RESULT 5  
 LOCUS DD212689 2773 bp DNA linear PAT 19-JAN-2006  
 DEFINITION NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS.  
 ACCESSION DD212689  
 VERSION DD212689.1 GI:85632167  
 KEYWORDS JP 2005518194-A/17.  
 SOURCE JP 2005518194-A/17.  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.  
 1 (bases 1 to 2773)

REFERENCE  
 AUTORS Moutilsen,S., Voldborg,B., Bratt,T., Nielsen,F.S. and Klysemer,S.  
 TITLE NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS  
 JOURNAL Patent: JP 2005518194-A 17 23-JUN-2005;

COMMENT  
 OS Artificial sequence  
 PN JP 2005518194-A/17  
 PD 23-JUN-2005

PF 15-NOV-2002 JP 2003544079  
 PR 16-NOV-2001 DK PA200101702.16-NOV-2001 US 60/331575 PT  
 soren moutilsen,bjoern voldborg,tomas bratt,finn scaustholm PT  
 nielsen,  
 PI steen klysemer  
 CC p220p2f expression vector for insect cells  
 FH Key Location/Qualifiers  
 FT misc\_recomb (561)..(566)  
 FT misc\_recomb /note='HindIII site'  
 FT misc\_recomb (573)..(578)  
 FT misc\_recomb /note='Aval site'  
 FT misc\_recomb (586)..(591)  
 FT misc\_recomb /note='EcoRI site'  
 FT misc\_recomb (593)..(598)  
 FT misc\_recomb /note='BamHI site'  
 FT misc\_recomb (625)..(630)  
 FT misc\_recomb /note='ClaI site'  
 FT misc\_recomb (629)..(634)  
 FT misc\_recomb /note='ClaI site'  
 FT misc\_recomb (1156)..(1161)  
 FT misc\_recomb /note='ApaLI site'  
 FT misc\_recomb (2128)..(2133)  
 FT misc\_recomb /note='PstI site'  
 FT misc\_recomb (2204)..(2209)  
 FT misc\_recomb /note='NcoI site'  
 FT misc\_recomb (2284)..(2289)  
 FT misc\_recomb /note='Aval site'  
 FT misc\_recomb (2294)..(2299)  
 FT misc\_recomb /note='Aval, SmaI, and XmaI site'  
 FT misc\_recomb (2551)..(2556)  
 FT misc\_recomb /note='ApaLI site'.

## FEATURES

source  
 1..2773  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"

## ORIGIN

Query Match 97.2%; Score 548; DB 2; Length 2773;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-180;  
 Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATGATGCTAATGCTTCAACAACAATCTGTTGAACGTG 60  
 DB 5 CATGATGATTAACAATGATGCTAATGCTTCAACAACAATCTGTTGAACGTG 64  
 QY 61 TTTTCATGTTTGCACACACCTTTTATCTCGTGCGCTCCCAACCAACTTTT 120  
 DB 65 TTTTCATGTTTGCACACACCTTTTATCTCGTGCGCTCCCAACCAACTTTT 124  
 QY 121 GCACGTGCAAAAAACAGCTTTTGCACGCGGCCCATATAGTACAACTTACGTTTC 180  
 DB 125 GCACGTGCAAAAAACAGCTTTTGCACGCGGCCCATATAGTACAACTTACGTTTC 184  
 QY 181 GTAGACATTTTATCATTAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 240  
 DB 185 GTAGACATTTTATCATTAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 244  
 QY 241 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTGCGGAGTCAAGTACGCGGCTTATC 300  
 DB 245 ATTGAACCTTTTTCAGTGCACAAAAAGTACGTGCGGAGTCAAGTACGCGGCTTATC 304  
 QY 301 GGGTCGCGTCTGTACGTAAGTACATTAATGACCGGACGAGTGTCTTATCGT 360  
 DB 305 GGGTCGCGTCTGTACGTAAGTACATTAATGACCGGACGAGTGTCTTATCGT 364  
 QY 361 GACAGACGCGCAGCTTCGTGTGTTGCTAACCGGACCGGACGCACTCCCTTATCGAACA 420  
 DB 365 GACAGACGCGCAGCTTCGTGTGTTGCTAACCGGACCGGACGCACTCCCTTATCGAACA 424  
 QY 421 GGAAGCGCTTCATATCAGCGCGGCTTATCTCATGCGCGTGAACGGAACGAGCGCC 480  
 DB 425 GGAAGCGCTTCATATCAGCGCGGCTTATCTCATGCGCGTGAACGGAACGAGCGCC 484

```

QY      481  GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACAGATTGAACAG 540
Db      485  GTCCCGCTTATCGCGCTATAAATACAGCCCGCAACGATCTGTAAACAGATTGAACAG 544
QY      541  CATCTGTT 548
Db      545  CATCTGTT 552

RESULT 6
AX766573      2773 bp      DNA      linear      PAT 25-JUN-2003
LOCUS      Sequence 60 from Patent WO03042244.
ACCESSION      AX766573
VERSION      AX766573.1 GI:32260450
KEYWORDS
SOURCE      .
ORGANISM      synthetic construct
              synthetic construct
              other sequences; artificial sequences.
REFERENCE      1
AUTHORS      Klysnier, S., Nielsen, F.S., Bratt, T., Voldborg, B. and Mouritsen, S.
TITLE      Novel immunogenic mimetics of multimer proteins
              Patent: WO 03042244-A 60 22-MAY-2003;
              Pharmexa A/S (DK) ; Klysnier, Steen (DK) ; Nielsen, Finn Stenusholm
              (DK) ; Bratt, Tomas (DK) ; Voldborg, Bjorn (DK) ; Mouritsen, Soren
              (DK)

FEATURES
    source      Location/Qualifiers
                1..2773
                /organism="synthetic construct"
                /mol_type="unassembled DNA"
                /db_xref="taxon:32630"
                /note="p2ZOP2F expression vector for insect cells"
    misc_recomb      561..566
                    /note="HindIII site"
    misc_recomb      573..578
                    /note="HindIII site"
    misc_recomb      586..591
                    /note="Aval site"
    misc_recomb      593..598
                    /note="EcoRI site"
    misc_recomb      625..630
                    /note="BamHI site"
    misc_recomb      629..634
                    /note="ClaI site"
    misc_recomb      1156..1161
                    /note="ClaI site"
    misc_recomb      2128..2133
                    /note="ApaLI site"
    misc_recomb      2204..2209
                    /note="PstI site"
    misc_recomb      2284..2289
                    /note="NcoI site"
    misc_recomb      2294..2299
                    /note="Aval site"
    misc_recomb      2551..2556
                    /note="Aval, SmaI, and XmaI site"
    misc_recomb      /note="ApaLI site"

ORIGIN
Query Match      97.2%; Score 548; DB 2; Length 2773;
Best Local Similarity 100.0%; Pred. No. 9,4e-180;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CAGGATGATTAACAAATGATGCTAATGCTTGAACAAATCTGTGAACATG 60
Db      5  CAGGATGATTAACAAATGATGCTAATGCTTGAACAAATCTGTGAACATG 64
QY      61  TTTTCATGTTGGCAACAGACCTTATATCTCGTGAGCTCCCAACCAACTTTT 120
Db      65  TTTTCATGTTGGCAACAGACCTTATATCTCGTGAGCTCCCAACCAACTTTT 124
QY      121 GCACTGCAAAAAAAGACGCTTTTGACGCGGGCCATACATAGTACAACTTACGTTTC 180

```

```

Db      125  GCATGCAAAAAAAGACGCTTTTGACGCGGGCCATACATAGTACAACTTACGTTTC 184
QY      181  GTAGACTATTTTACATAATAGCTTACACCGTTGTATAGCTTCCAATACATACACAC 240
Db      185  GTAGACTATTTTACATAATAGCTTACACCGTTGTATAGCTTCCAATACATACACAC 244
QY      241  ATGAACCTTTTGCAGTGCAAAAAAGTACGTCGCGAGTCACGTAGGCGGCTTATC 300
Db      245  ATGAACCTTTTGCAGTGCAAAAAAGTACGTCGCGAGTCACGTAGGCGGCTTATC 304
QY      301  GGGTCGCGTCTCTGTACATACGATCATATTCGAGCCGAGAGTGTCTTATCGT 360
Db      305  GGGTCGCGTCTCTGTACATACGATCATATTCGAGCCGAGAGTGTGTCTTATCGT 364
QY      361  GATAGAGCGCCAGCTTCTGTGTGTGTACACCGGACCGGACCGAATCTCTTATCGGAACA 420
Db      365  GATAGAGCGCCAGCTTCTGTGTGTGTACACCGGACCGGACCGAATCTCTTATCGGAACA 424
QY      421  GGACGCGCTCCATATCAGCGCGCGCTTATCTCATGCGGTGACCGGACGAGCGGCC 480
Db      425  GGACGCGCTCCATATCAGCGCGCGCTTATCTCATGCGGTGACCGGACGAGCGGCC 484
QY      481  GTCCGCTTATCGCGCTTATAATATACAGCCCGCAACGATCTGTGAACAGATTGAACAG 540
Db      485  GTCCGCTTATCGCGCTTATAATATACAGCCCGCAACGATCTGTGAACAGATTGAACAG 544
QY      541  CATCTGTT 548
Db      545  CATCTGTT 552

RESULT 7
AB175497/c
LOCUS      2284 bp      DNA      linear      VRL 29-MAR-2005
DEFINITION      Hyphantria cunea nucleopolyhedrovirus ie2, hycu-ep32, odv-e56 gene
                  for immediate early gene 2, early protein 32, occlusion derived
                  virus envelope protein 56, partial and complete cde.
ACCESSION      AB175497
VERSION      AB175497.1 GI:62086138
KEYWORDS
SOURCE      .
ORGANISM      Hyphantria cunea nucleopolyhedrovirus
                  Hyphantria cunea nucleopolyhedrovirus
                  Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
                  Nucleopolyhedrovirus.
REFERENCE      1
AUTHORS      Shirata, N., Ikeda, M. and Kobayashi, M.
TITLE      Hyphantria cunea nucleopolyhedrovirus (NPV) ep32 gene is involved
                  in the restriction of Bombyx mori NPV replication in B. mori cells
                  Unpublished
JOURNAL      2 (bases 1 to 2284)
REFERENCE      Shirata, N., Ikeda, M. and Kobayashi, M.
AUTHORS      Direct Submission
TITLE      Submitted (26-MAR-2004) Michihiro Kobayashi, Nagoya University,
                  Graduate School of Bioagricultural Sciences, Laboratory of
                  Biodynamics, Chikusa, Nagoya, Aichi, 464-8601, Japan
                  (E-mail: michihiro@agr.nagoya-u.ac.jp, Tel:81-52-789-4038,
                  Fax:81-52-789-4038)

FEATURES
    source      Location/Qualifiers
                1..2284
                /organism="Hyphantria cunea nucleopolyhedrovirus"
                /mol_type="genomic DNA"
                /strain="N9"
                /db_xref="taxon:28288"
                complement(1..252)
                /gene="ie2"
                complement(1..252)
                /codon_start=1
                /product="immediate early gene 2"
                /protein_id="BAD91566.1"
                /db_xref="GI:62086139"
                /translation="MSRINNADTPTNRRNLISARGRRRLTSPSTPMPPTPRTAPAMP
                TPRRSLTPRHVGVDRNAPLRASYTINNSRYNVHGDADF"

```

```

gene
  817. .1755
  /gene="hycu-ep32"
  817. .1755
  /gene="hycu-ep32"
  /codon_start=1
  /product="early protein 32"
  /protein_id="BAD91567.1"
  /db_xref="GI:62086140"
  /translation="MKKQQQLSAGQQRAANSARLTIKKQADNVKLSDSMSAWLAT
  APTBSTPLVPLNSMNTTGANVQFPPVQTAHINIPIATATPFTVATRCMGN
  IDCKFACNPKCADLETGTINNAYFKQDEAKGVTIQLQOQDKEIDQFSA
  DNNSCNLSIFKIDMLFKQKENGFNDRRRACALNNQNDIVLKYKEDFDDDTQNI
  SLKMLSVTSGNFNIIFFPETNYSWKKLASNNLITVVDNQDETVIQTFNKSVEPFC
  VALKQIQLITMSQLYGRVEINNNMAMTFNMDLIN"
  1877. .>2284
  /gene="odv-e56"
  1877. .>2284
  /gene="odv-e56"
  /product="odv-e56"
  /codon_start=1
  /product="occlusion derived virus envelope protein 56"
  /protein_id="BAD91568.1"
  /db_xref="GI:62086141"
  /translation="MSFSPSLRRVNRVYPSNASFVTDNSRLTSTPAGTNNVLAPEST
  RNLGNGRYEPGYLSNNQFISAGDINRITRSNDVPRIRNVFOGINDPQINSQLRRA
  DNPDAGFHYKRTRGDAVRQNFPEPTVRSPEGVD"

```

## ORIGIN

```

Query Match      24.5%; Score 138.4; DB 10; Length 2284;
Best Local Similarity 78.3%; Pred. No. 2.5e-36;
Matches 166; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

```

QY 81 CACCTTATCTCGTGCCCTCCACCACCAACTTTTTCGATCGAAAAACAGCCT 140
   |||||
DB 686 CACCGCTTGATATCTGCTGTGCACACGAACTTTTTCAGTGCAAAAACAGCTT 627
   |||||
QY 141 TTTCACGCGGGCCCATACATAGTACAACTCTACGTTGCTAGACTATTTTACATAAT 200
   |||||
DB 626 ATGAACACAGCCCCATATACAGGCAAGCTATCTATTTGCTAGACTATTTTACATAAT 567
   |||||
QY 201 AGCTTACACCGCTGTATACGCTCCAAATACATACACACATTGAACCTTTTTCAGTGC 260
   |||||
DB 566 AGCTTACACGCTGTATACGCTCCAAATACATACACACATTGAACCTTTTTCAGTGC 507
   |||||
QY 261 AAAAAGTAGCTGTGCGCAGTCACTAGGCGC 292
   |||||
DB 506 AAAAATACGTGTGCGAGTCACTACTCGC 475
   |||||

```

```

RESULT 8
AP009046/c      132959 bp      DNA      circular VRL 31-JAN-2006
LOCUS DEFINITION Hyphandria cunea nucleopolyhedrovirus genomic DNA, complete genome.
AP009046
VERSION AP009046.1 GI:86198125
KEYWORDS
SOURCE
  Hyphandria cunea nucleopolyhedrovirus
  Hyphandria cunea nucleopolyhedrovirus
  Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
  Nucleopolyhedrovirus.
ORGANISM

```

```

REFERENCE
  1 Felipealves, C.A., Ikeda, M. and Kobayashi, M.
  Identification and characterization of Hyphandria cunea
  nucleopolyhedrovirus homologous repeated regions
  Virus Gene 25, 281-290 (2002)
  2 Ikeda, M., Yanagimoto, K. and Kobayashi, M.
  Identification and functional analysis of Hyphandria cunea
  nucleopolyhedrovirus lap genes
  virology 321 (2), 359-371 (2004)
  3 15051395
  Ikeda, M., Shikata, M., Shirata, N., Chaeychomari, S. and Kobayashi, M.
  Gene organization and sequencing of the Hyphandria cunea

```

```

JOURNAL
REFERENCE
  4 (bases 1 to 132959)
  Ikeda, M. and Kobayashi, M.
  Direct Submission
  Submitted (17-AUG-2005) Motoko Ikeda, Nagoya University, Graduate
  School of Biological Sciences; Chikusa, Nagoya, Aichi
  464-8601, Japan (E-mail: mochikoeag@nagoya-u.ac.jp,
  tel: 81-52-789-4038, fax: 81-52-789-4036)
  Location/Qualifiers
  1. 132959
  /organism="Hyphandria cunea nucleopolyhedrovirus"
  /mol_type="genomic DNA"
  /db_xref="taxon:28288"
  1. .738
  /gene="polh"
  /locus_tag="HynVgp001"
  1. .738
  /gene="polh"
  /locus_tag="HynVgp001"
  /note="orf1 polh"
  /codon_start=1
  /product="polyhedrin"
  /protein_id="BAE72290.1"
  /db_xref="GI:86198126"
  /translation="MPDFSTRPTIGRTTYDNTKYKSLGSLVTKAKKKHLEHDE
  KFLDPLDHYWVAEDPFLGREGKQKTLFKETIRNVKPDPTMKLVNMSGKEFLRETWRP
  VEDSFPIVNDQEVDFLVLNLRPTPNRCYKLAQHALRMNDVNPVHEVIRLVEBSY
  VGNMNNYRSLSIARCGGCPIMNIHSEYTSFEQFVNRVIMENFYKPIVYIGTDSGEE
  ELLIEVSLVKEVFEAPDAPLFGPAY"
  complement (735. .2354)
  /gene="cap"
  /locus_tag="HynVgp002"
  complement (735. .2354)
  /gene="cap"
  /locus_tag="HynVgp002"
  /note="orf2 1629capsid"
  /codon_start=1
  /product="1629capsid"
  /protein_id="BAE72291.1"
  /db_xref="GI:86198127"
  /translation="MMEROQVSRSYLINNELDAIDAGTFLGRVAGBEAHDLRKCLND
  LTRIDRTAVLNILKLAENIYTDVAHYMDRVDASHSATLGRMRLMTYDOLNRR
  NNDITLRBGLRLDTYHDAVEGUISDPFEYSKYTSQQKRVRTSLFNLEQQQQQ
  QQQQQQQQQQQPFAQQQPTPTQLPATTTTISRPTNEEFNYTANSEDLVOTPI
  KEPPVPKPPHLTRPASIDTNVPEPPPTPPPPPMPPPPPPPPPPPPPPPPPLSL
  PPIDNLLNAIVSEPMKAGDARDALBLEQISGTLKVEPPAAPVSDARNLELQ
  ITVGRLLKKTIRPKGKEKPPPPSNDVAVILANTLRRRRKMDTDSASGISSESDSM
  VPDIEKADKRELKTYVNLHNFPSKSNAPKTDNDIELRLVDTAATSLDCKPTTENV
  EKANVTLNSFRERVVVPANALDSQAPALYAKDAOFYVAIBDLILAGKTEBARHLK
  LVNAPPDAVSELSITYANQLISANQLIASSVA"
  2353. .3177
  /gene="pk1"
  /locus_tag="HynVgp003"
  2353. .3177
  /gene="pk1"
  /locus_tag="HynVgp003"
  /note="orf3 pk1"
  /codon_start=1
  /product="Protein kinase 1"
  /protein_id="BAE72292.1"
  /db_xref="GI:86198128"
  /translation="MDATLQSAVQFVKDCVLTLPKYVNGRFGKIDVLAHRPTSKLYLR
  KTSANFNADENIVHDLADHPNFIMNFCVGTPTAAAIWVDYVQCPDLFETLQTEG
  ALENALVSNIVROLCDALNDLHAATGYIHNDVLALEVLVLFKARDRYVLCGYCKKEH
  SGVHDGTLVEYSPKIRRHNYARSPDVAVGVTYKLLTGKGHPERSGVDEVLDS
  MKRRQGYNDIATLKSRNNAHARDPFCYCLTFKFLDGRLLIDYKRIVGSPLVTGNTY"
  complement (3168. .3413)
  /locus_tag="HynVgp004"
  complement (3168. .3413)
  /locus_tag="HynVgp004"
  /note="orf4"
  /codon_start=1

```

```

gene
  1. .738
  /gene="polh"
  /locus_tag="HynVgp001"
  1. .738
  /gene="polh"
  /locus_tag="HynVgp001"
  /note="orf1 polh"
  /codon_start=1
  /product="polyhedrin"
  /protein_id="BAE72290.1"
  /db_xref="GI:86198126"
  /translation="MPDFSTRPTIGRTTYDNTKYKSLGSLVTKAKKKHLEHDE
  KFLDPLDHYWVAEDPFLGREGKQKTLFKETIRNVKPDPTMKLVNMSGKEFLRETWRP
  VEDSFPIVNDQEVDFLVLNLRPTPNRCYKLAQHALRMNDVNPVHEVIRLVEBSY
  VGNMNNYRSLSIARCGGCPIMNIHSEYTSFEQFVNRVIMENFYKPIVYIGTDSGEE
  ELLIEVSLVKEVFEAPDAPLFGPAY"
  complement (735. .2354)
  /gene="cap"
  /locus_tag="HynVgp002"
  complement (735. .2354)
  /gene="cap"
  /locus_tag="HynVgp002"
  /note="orf2 1629capsid"
  /codon_start=1
  /product="1629capsid"
  /protein_id="BAE72291.1"
  /db_xref="GI:86198127"
  /translation="MMEROQVSRSYLINNELDAIDAGTFLGRVAGBEAHDLRKCLND
  LTRIDRTAVLNILKLAENIYTDVAHYMDRVDASHSATLGRMRLMTYDOLNRR
  NNDITLRBGLRLDTYHDAVEGUISDPFEYSKYTSQQKRVRTSLFNLEQQQQQ
  QQQQQQQQQQQPFAQQQPTPTQLPATTTTISRPTNEEFNYTANSEDLVOTPI
  KEPPVPKPPHLTRPASIDTNVPEPPPTPPPPPMPPPPPPPPPPPPPPPPPLSL
  PPIDNLLNAIVSEPMKAGDARDALBLEQISGTLKVEPPAAPVSDARNLELQ
  ITVGRLLKKTIRPKGKEKPPPPSNDVAVILANTLRRRRKMDTDSASGISSESDSM
  VPDIEKADKRELKTYVNLHNFPSKSNAPKTDNDIELRLVDTAATSLDCKPTTENV
  EKANVTLNSFRERVVVPANALDSQAPALYAKDAOFYVAIBDLILAGKTEBARHLK
  LVNAPPDAVSELSITYANQLISANQLIASSVA"
  2353. .3177
  /gene="pk1"
  /locus_tag="HynVgp003"
  2353. .3177
  /gene="pk1"
  /locus_tag="HynVgp003"
  /note="orf3 pk1"
  /codon_start=1
  /product="Protein kinase 1"
  /protein_id="BAE72292.1"
  /db_xref="GI:86198128"
  /translation="MDATLQSAVQFVKDCVLTLPKYVNGRFGKIDVLAHRPTSKLYLR
  KTSANFNADENIVHDLADHPNFIMNFCVGTPTAAAIWVDYVQCPDLFETLQTEG
  ALENALVSNIVROLCDALNDLHAATGYIHNDVLALEVLVLFKARDRYVLCGYCKKEH
  SGVHDGTLVEYSPKIRRHNYARSPDVAVGVTYKLLTGKGHPERSGVDEVLDS
  MKRRQGYNDIATLKSRNNAHARDPFCYCLTFKFLDGRLLIDYKRIVGSPLVTGNTY"
  complement (3168. .3413)
  /locus_tag="HynVgp004"
  complement (3168. .3413)
  /locus_tag="HynVgp004"
  /note="orf4"
  /codon_start=1

```

/product="ORF4 peptide"  
/protein\_id="BAE72293.1"  
/db\_xref="GI:86198129"  
/translation="MGDVNKNLLIEHNAFLVHPVTALTPDDNTFDISGNVATPDKCLN  
VKATPEFLNSATCDNVAALNNQPAETLSNOCNIQI"  
complement(3419..4501)  
/gene="pe38"  
/locus\_tag="HynVgp005"  
complement(3419..4501)  
/gene="pe38"  
/locus\_tag="HynVgp005"  
/note="orf5 pe38"  
/codon\_start=1  
/product="pe38"  
/protein\_id="BAE72294.1"  
/db\_xref="GI:86198130"  
/translation="MSNRHYHYSFNSPTORRVOEOMLKTLLNKPVEFRATCAVCLD  
TYKQSHNITDFLMSNCTHAFCYCVLNNMYCNTMNI PRATIDCPMKVNTWRSCE  
PNAVSCKFIKKTGDRVPCQGFVEAVQIKORCAADDEPAVSQNNKLIQOETER  
LOKTVDKHKOAEERLQRIYDKHKOAEERLOETVNVKROGEBRLOEITINCKROHDIAW  
SSSECVYLAONKIDLOAQDLKSKITQTLADQNRTPVALFKYLQNFVNHENQHHH  
NDENDEHGHQLEIKKNVAVNVEFENNAKQTNLHQFNSRIVSTVADMMIIEBKIKS  
LQTHVFGIACLPQSVNINVSLLPFDDE"  
4710..5602  
/standard\_name="hr1"  
complement(5764..6891)  
/gene="ie2"  
/locus\_tag="HynVgp006"  
complement(5764..6891)  
/gene="ie2"  
/locus\_tag="HynVgp006"  
/note="orf6 ie2"  
/codon\_start=1  
/product="ie2"  
/protein\_id="BAE72295.1"  
/db\_xref="GI:86198131"  
/translation="MSRIINNADPTNRNRLSLARGRRRLTYSPESTPMPTQGTAPAP  
TPBORSLTPEHVYDGRNAPLAGYTNINSRYVNGPAEFNPEDDDSIIFPDHAE  
QTORAVNLHESLTTTPRSPDYPHSPGQVUIDSDYNDDBRMELOILLHAEER  
QTRQAPAREPQEDVEVLCHICSGFTDVKNNNSNVITSECHAVCFCTYSIYNPKKA  
YKSCICRRITLTCAVNRAGVELSTRVARDNKLIKQHWMLTESMPPNRDQITIE  
ELDLADLRATTAARAHEVNMIKSDMLLQQQVDYFNLELQOELNKNVQLQKNDTL  
SAANTFLQNDLDAQVAESKIKMQFVQHEHFLKFKFSYVW"  
7249..7434  
/locus\_tag="HynVgp007"  
7249..7434  
/locus\_tag="HynVgp007"  
/note="orf7"  
/codon\_start=1  
/product="ORF7 peptide"  
/protein\_id="BAE72296.1"  
/db\_xref="GI:86198132"  
/translation="MGLCSITCFPALOKSSCVTSKIIKWRREGAGKDKKWKYTRGSK  
FENTVQNLCCDSNQLVS"  
7456..8394  
/gene="ep32"  
/locus\_tag="HynVgp008"  
7456..8394  
/gene="ep32"  
/locus\_tag="HynVgp008"  
/note="orf8 ep32"  
/codon\_start=1  
/product="ep32"  
/protein\_id="BAE72297.1"  
/db\_xref="GI:86198133"  
/translation="MKQQQLISAQQQRAANSARRLTIKKORADWVKKLSDSMSAMLAT  
APRPSYPLVFLNMSMNTTAGANVQFQVIVOTAHIRIPITATVRFVATHCRCMHN  
IDGHKFCAPNCRKADLETGNTDINNAFYKLQEDAEKGLVTIEQLQQQDKEIQYFSA  
DNNSCNLSIKRIDNLFKQKEYNGFNDRBRRTATYNNQNDIVLYKEDFPDDDTQNTI  
SLKMLSVTSGNRNILLPFTNNASMKKLAENNLITVYDNOQETPVIOTFKKSVBPPVC  
VALKQIKQITMSQLYQREVEINNNNAATFENFMDLIN"  
8516..9646  
/gene="odv-e56"

CDS  
/locus\_tag="HynVgp009"  
8516..\_9646  
/gene="odv-e56"  
/locus\_tag="HynVgp009"  
/note="orf9 odv-e56"  
/codon\_start=1  
/product="ODV-E56"  
/protein\_id="BAE72298.1"  
/db\_xref="GI:86198134"  
/translation="MSFNSLRVNRVYPSNMAFVTJDNRLTSTPAGFTNVLANPST  
RLNGNRYEBGYLLSNNOFIISAGDINRITRSNDVPRIRNVFQGINDPQINSLNQLERA  
DNVPDAGFHYKTRGDVNRQNPETNRGSPGVNDLQONPRILNMLKQTAGVVL  
LACGAYLASAATLVODIICALNTEGSGSYFVSGSDGESAESCLLNTRCRRPNNAT  
SEVAICANDPLISDTAELOAICNGFVNEREQYTCRQSDPADPDSQFVDSIDLPQO  
TINCIRPNYLAIDLIGDLGHDLGQDGLGKLSNSSDSIGNTLMPITIIIGANVFLGL  
VYLIRFLTKSGGGSNSVANSNPSTVMPQPLQTRG"  
complement(9807..11489)  
/gene="ie1"  
/locus\_tag="HynVgp010"  
complement(9807..11489)  
/gene="ie1"  
/locus\_tag="HynVgp010"  
/note="orf10 ie1"  
/codon\_start=1  
/product="IE1"  
/protein\_id="BAE72299.1"  
/db\_xref="GI:86198135"  
/translation="MKRMAALHRSNFTGPTPSHTLFNTATLPLNULINCTYNAV  
APDFAFVFNETSMTLLEQTNLTHVDVNFELARSTASAFVTGRANSLEEVOLV  
EQQKGRKRSLELSDSDSDSKGKTLVNRKMRQRYKATIQNKTLTEBOHST  
EICTVAPSEQIAKYFSGDSVYLDKINSQGSANFSDHISBTGYVFIYKKSFEKP

Query Match 24.5%; Score 138.4; DB 10; Length 132959;  
Best Local Similarity 78.3%; Pred. No. 3.3e-36;  
Matches 166; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
Qy 81 CACCTTATCTAGTGGTGCCCTCCCAACCACTTTTGGACATGCAAAAAACAGCT 140  
Db 7325 CACCGCTGATATATCTTCTGCTGTGCACAGATGAAGCTTTTGGCGTCAAAAAACAGCT 7266  
Qy 141 TTTCACGCGGCGCCCATCATCATGACAACTGATGCTTGTGACTATTTTACATTAAT 200  
Db 7265 ATGGAACACAGCCCAATATACAGAGCAAGATCTTCTGATGACTATTTTACTGGAAT 7206  
Qy 201 AGCTCACCGCTGTATAGCTCCCAATACATGACACACATGAACCTTTTGGACATGC 260  
Db 7205 AGTCTACACTGTCTATACGCTCTATATACATGACAGAGTAACTTTTGGCAGTGC 7146  
Qy 261 AAAAAAGTACGTGTGCGACGTCACTGAGCGC 292  
Db 7145 AAAAAATACGTGTGAGAGTCACTGACTCGC 7114  
RESULT 9  
AY043265/c 118584 bp DNA circular VRL 19-AUG-2001  
LOCUS DEFINITION  
AY043265 Epiphyas postvittana nucleopolyhedrovirus, complete genome.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Epiphyas postvittana NPV  
Epiphyas postvittana NPV  
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
Nucleopolyhedrovirus.  
REFERENCE  
1 (bases 1 to 118584)  
Hynk,O., Dellow,R.A., Olsen,M., Caradoc-Davies,K.M.B., Drake,K.  
and Ward,V.K.  
TITLE  
The complete sequence of the Epiphyas postvittana  
nucleopolyhedrovirus genome  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 118584)  
Hynk,O., Dellow,R.A., Olsen,M., Caradoc-Davies,K.M.B., Drake,K.  
and Ward,V.K.

```

TITLE      Direct Submission
JOURNAL    Submitted (03-JUL-2001) Microbiology, University of Otago, PO Box
           56, Dunedin, New Zealand
FEATURES   Location/Qualifiers
           source      1. 118584
                       /organism="Epiphyas postvittana NPV"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:70600"
                       /note="EppomNPV"
                       join(117505..118584,1..609)
                       /note="ORF136"
CDS         /codon_start=1
           /product="ORF1629"
           /protein_id="AAK85700.1"
           /db_xref="GI:15213261"
           /translation="MDRQYQSVKSYLIKNNANFADANAFLLTLIAGPEAHNVKONLIGG
           TVLRNRAVADLLKLAENIYNDTAYMEVDVNDLRHVTTLMRKQMLISTDQVRRLT
           LANIVSRVHLLRLEYIYNDVEITTLISGDFYEESYKARQYALSIQWPPPPVITPLP
           PPLTPHQVPPPPPLTPILOGLIPVADVATPSPSTLSKSTTLDSREYFSNAMSQ
           LPVTPKPLIPVKEHKKPKKRTIISLPMKATNIDDKLPAPPPPIPPSPPL
           PSNDLPPPPPLPSGNVPPPPPIEGMLDMLINALIAGNNSEKVTQARDQMLNL
           IKKGVTLKPSKTKNSDKKVDRLDNLNSIKIGVLAKEIKVNTQPLPEAVTDISYIA
           GTLKRRANISQSESSSSNSQSMDEGNTQIRKANQDLKXAVLVNFPATTHAY
           RTNSELPKLDNYFSLIDRKPRSVENVEANKIIDLKERVKLTSNQLDNAAQSLYI
           NDNOFYIOVEDLIIFAGRYADAKMHLDLATTESGDBRLRLKKFANDDAVVY"
           complement(606..1343)
gene        /gene="pol"
           /complement(606..1343)
CDS         /product="unknown"
           /protein_id="AAK85566.1"
           /db_xref="GI:15213126"
           /translation="MPDYSYRPTIGRTYVYDNKKYKLAGSVIKNAKKKHLLEHED
           KEDPLDHYVAADPLVPGNKQKLTLEKIRVVKPDTMKLIYNSGKEFLREMTWRP
           VHSDFPIVNDQEVMDVFLVNLNLRPTRNRCYKLAQALHMDQDYVPEHVRILVSESY
           VGMNRYRISLAKGGGCPIMNIHSETNSFESFARVIVENTYKPIVYIGTDSGE
           ELIEVSLVFKEKFAFDAPLFTGPAL"
           complement(1434..1691)
CDS         /note="ORF2"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85566.1"
           /db_xref="GI:15213127"
           /translation="MQFLREIYDNYLQRAATIMCGTCSETNANVOYREHLLTYHKK
           CPTQLSCVRCDDQSSWGFSGALMEHMYFCLQNYQNGV"
           complement(1713..2090)
CDS         /note="ORP3"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85567.1"
           /db_xref="GI:15213128"
           /translation="MGASODSEBQKILKLFYCWSSQTAGLNDNDKDLHCLCYDLERFVGA
           HLMROADEKAKKKKSAKALKRVEILAADHMLAATPACTPDGGRNSTLSQQLDDI
           ARKQDLYDRILYHQLKODRLIKQ"
           complement(2104..2715)
gene        /gene="lef-2"
           /complement(2104..2715)
CDS         /gene="lef-2"
           /note="ORP4"
           /codon_start=1
           /product="LEF-2"
           /protein_id="AAK85568.1"
           /db_xref="GI:15213129"
           /translation="MDNVAVMNSPANTATIKKNETYLIDPNDFVLTLSPTVPRRG
           OLVEISGMRLFALLTMSKPEPKOPARBSKRNICAKADSSQTLHKSLTGKILCLIC
           IAKIMAEISNASVPRGMKPRPENCYLGAVLTCTKCSACILGALLHFYKADACV
           GRTTHLLKSQDYKPSNCAKMSMKLCPOANNCKGLNPICNF"
           complement(2717..3070)
CDS         /note="ORP5"
           /codon_start=1
           /product="unknown"

CDS         /protein_id="AAK85569.1"
           /db_xref="GI:15213130"
           /translation="MILRYRIYIFVFPRESRTSKKPRISHIKLFFPYATLHYNNMR
           INNAFVAVAGHDYDRQDLKRDINLSHVSHELCKRSTGTGDCNRLLSESTDYKAPTV
           IKTTAAEOHTICDKL"
           complement(2990..3424)
CDS         /note="ORF6"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85570.1"
           /db_xref="GI:15213131"
           /translation="MADARELPRSTLTTEYKENVVAQIDHLNLBALINAKTTIANV
           BRGFPVDRNAFVAACMDVIVHTYASSDDKTTLQPEQYFPCFNEIADVPDDHS
           ITRYLCAQGTCLVDNPLNVFTEDGVQRMELQRIINAGEP"
           complement(3426..3935)
gene        /gene="pcp-1"
           /complement(3426..3935)
CDS         /note="pcp-1"
           /codon_start=1
           /product="PCP-1"
           /protein_id="AAK85571.1"
           /db_xref="GI:15213132"
           /translation="MEPDRMHDYTCGVNIEGNTFICPKVPLREEMFVYTNLDNRT
           IKRLIQQCSLGAVIDLTPTFRYTDANNRDBELIKKIRVPOQVLPDENVOQLFPDV
           VKHFTARCPGMLIGVCHTGLNRTGYLVCRMINIINISPDQALARETARGHIERQ
           NYIQLHLTI"
           complement(4168..5145)
CDS         /note="ORP8"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85572.1"
           /db_xref="GI:15213133"
           /translation="MSLSKLLVYDYYSNMGCVNBOYGASVHLKYIVDEHITNSYVG
           DYPWMEIBDTRQILSIALCVNDAQCHMLDNLNTERLATWKCQGTQKVCHEVYVL
           KALDALVPIIDKSGSGFQVNDVDPDTLLITKLPVLAQVLRPYVQKLEHVTQVFN
           PMSDTPGVGNKFCVUTVTRIRIINNSVPABELTRLOKAVETIOPNAGKNCALAEVMA
           NVYRGCGIGKHQFABHKTANNHIMFKSRGGQTTWDEKYSFSVTKDSFRICKETYL
           DLKSHVDSLVIYNTCTDKQNAFMDLCCQSEBDLDVDCYDYIVNMLFFK"
           complement(5163..5516)
CDS         /note="ORF9; unique to EppomNPV"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85573.1"
           /db_xref="GI:15213134"
           /translation="MSSNSFMDVDNTHENVETNVDTLKPDINWNSDSSQARCEK
           GAPPVFAACVPLESPPLNGQYKIDPTLPALQKITTDYANNTPVLASVKPNNYQDKN
           TVAAAYNNVLKNNIT"
           complement(5531..6490)
CDS         /note="ORF10"
           /codon_start=1
           /product="unknown"
           /protein_id="AAK85574.1"
           /db_xref="GI:15213135"
           /translation="WFDWFGWMSFVEHQIQHPLANHFNNADYKOAYAVDRQAQSLV
           NRSVFKCHPTFRFRVYIDDDNGRCRRVDFCKGLEINHELMNCMDSKVRAHLNEI
           VKTTPVEIISPDMSGTVYATKGLIOILQQLSFYEYQDVLLAIKTDKGYDCDVRDNI
           KTVLKIKITLVNNSDKFINAKKLFENQVCARFEQLEORLETLEKRVADATPMGVIIPR
           DVNKHQHLAVFVNOERGNTQIGFARGQSEYFPKRLLEPFEEDMHKLETVHNPQWAV
           QCIKDRFTISNGYKIKKVSRRBRVIVDNCINAAKOIVNNVIIS"
           complement(6427..7182)
gene        /gene="lef-1"
           /complement(6427..7182)
CDS         /gene="lef-1"
           /note="ORF11"
           /codon_start=1
           /product="LEF-1"
           /protein_id="AAK85575.1"
           /db_xref="GI:15213136"
           /translation="MATYVYTLDRVNLNMSIAYNNSRKFAFMTLKQWIVHVRHFDN
           ALQNLNFINSSVDVHVYPLDEGGRVYIDAPFKNYVKTSLMKINVGATALLFF
           TEMHVRHIMFSGNRGPHMLIKRADGPKIGALSVREHWYRIPEKPTRLNMRDTRNSF
           AHCMQRAVNMVQDQLQSSKPDLEILHFIPVDVQVPCNPTQIIRAPSPSYNHKGTQFS

```



gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"  
8914. .9528  
/note="ORF13"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85577.1"  
/db\_xref="GI:15213138"  
/translation="METPRLPSSFPATPRRAALGTVKTVTTTVSGRGLTRDEBNR  
IFOVAQLRKTRLSFTKLTQLOKRVYNNQRLVRKNNIADLAQLERRRCRSKSK  
YFAVLCNGVLLTISGSGFVRQVRVNNCAVGEQIFCRRNDCAARDQLIAEALAS  
IGSDVITSANNRPKILNKVYNAKFIYQGLHGFNNYPNTH"  
9497. 10123  
/note="ORF14"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85578.1"  
/db\_xref="GI:15213139"  
/translation="MDLTITLPLISLKNIEBPGRSECFKLTSKREDAEFCLVNCKSP

Query Match  
Best Local Similarity 19.4%; Score 111.8; DB 10; Length 118584;  
Matches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"  
8914. .9528  
/note="ORF13"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85577.1"  
/db\_xref="GI:15213138"  
/translation="METPRLPSSFPATPRRAALGTVKTVTTTVSGRGLTRDEBNR  
IFOVAQLRKTRLSFTKLTQLOKRVYNNQRLVRKNNIADLAQLERRRCRSKSK  
YFAVLCNGVLLTISGSGFVRQVRVNNCAVGEQIFCRRNDCAARDQLIAEALAS  
IGSDVITSANNRPKILNKVYNAKFIYQGLHGFNNYPNTH"  
9497. 10123  
/note="ORF14"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85578.1"  
/db\_xref="GI:15213139"  
/translation="MDLTITLPLISLKNIEBPGRSECFKLTSKREDAEFCLVNCKSP

Query Match  
Best Local Similarity 19.4%; Score 111.8; DB 10; Length 118584;  
Matches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"  
8914. .9528  
/note="ORF13"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85577.1"  
/db\_xref="GI:15213138"  
/translation="METPRLPSSFPATPRRAALGTVKTVTTTVSGRGLTRDEBNR  
IFOVAQLRKTRLSFTKLTQLOKRVYNNQRLVRKNNIADLAQLERRRCRSKSK  
YFAVLCNGVLLTISGSGFVRQVRVNNCAVGEQIFCRRNDCAARDQLIAEALAS  
IGSDVITSANNRPKILNKVYNAKFIYQGLHGFNNYPNTH"  
9497. 10123  
/note="ORF14"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85578.1"  
/db\_xref="GI:15213139"  
/translation="MDLTITLPLISLKNIEBPGRSECFKLTSKREDAEFCLVNCKSP

Query Match  
Best Local Similarity 19.4%; Score 111.8; DB 10; Length 118584;  
Matches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"  
8914. .9528  
/note="ORF13"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85577.1"  
/db\_xref="GI:15213138"  
/translation="METPRLPSSFPATPRRAALGTVKTVTTTVSGRGLTRDEBNR  
IFOVAQLRKTRLSFTKLTQLOKRVYNNQRLVRKNNIADLAQLERRRCRSKSK  
YFAVLCNGVLLTISGSGFVRQVRVNNCAVGEQIFCRRNDCAARDQLIAEALAS  
IGSDVITSANNRPKILNKVYNAKFIYQGLHGFNNYPNTH"  
9497. 10123  
/note="ORF14"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85578.1"  
/db\_xref="GI:15213139"  
/translation="MDLTITLPLISLKNIEBPGRSECFKLTSKREDAEFCLVNCKSP

Query Match  
Best Local Similarity 19.4%; Score 111.8; DB 10; Length 118584;  
Matches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"  
8914. .9528  
/note="ORF13"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85577.1"  
/db\_xref="GI:15213138"  
/translation="METPRLPSSFPATPRRAALGTVKTVTTTVSGRGLTRDEBNR  
IFOVAQLRKTRLSFTKLTQLOKRVYNNQRLVRKNNIADLAQLERRRCRSKSK  
YFAVLCNGVLLTISGSGFVRQVRVNNCAVGEQIFCRRNDCAARDQLIAEALAS  
IGSDVITSANNRPKILNKVYNAKFIYQGLHGFNNYPNTH"  
9497. 10123  
/note="ORF14"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAK85578.1"  
/db\_xref="GI:15213139"  
/translation="MDLTITLPLISLKNIEBPGRSECFKLTSKREDAEFCLVNCKSP

Query Match  
Best Local Similarity 19.4%; Score 111.8; DB 10; Length 118584;  
Matches 131; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

gene  
7254. .8732  
/gene="egt"  
7254. .8732  
/gene="egt"  
/note="ORF12"  
/product="Egt"  
/protein\_id="AAK85576.1"  
/db\_xref="GI:15213137"  
/translation="MHFAVLLTLTAQAQAAHAIILAVLPTPAVSHLVYKRAYVOALANK  
CHNVYIKPOLNLYDSSNAHMGCHIEIQIADSSBQYKYLVSAGGAFRRGVADDTQ  
VTANNVGLIDHFQKDFNNVVOQELLCRTEDVVAEAPADYALVFGHLPRAPIQ  
IAPGVLDENFDPAGAVARHPRKYPNIWSTFVAGRAGVLSERVLNBEELLARVSDK  
LKIIGPPTPTIRLELNVOQLNLNHPVDDNRVLPSPVOYLGGGLHQTQPNL  
OKTLEQLNASVNGIYVFESSIDTSSIHDEFLQNLVTFGLNRRVYLMKYDAAV  
DSIKLPPNVITONFQRAVLHKKNTVAFTQSGLOSSEALHAQVPMCLPMWGOE  
HNAKIQEFKVAARTLNTVVSASQILATIDVYVKKITQTRAEILRAVIDDEIAPA  
DAIKTERVYIKFHGDITTPARSLKSPANLDSNFIISFPL"

TITLE	COMPARATIVE ANALYSIS OF THE GENOMES OF RACHIDPLUSIA OU AND AUTOGRAPHA CALIFORNICA MULTIPLE NUCLEOPOLYDROVIRUSES
JOURNAL	J. Gen. Virol. 84 (Pt 7), 1827-1842 (2003)
PUBMED	12810877
REFERENCE	3 (bases 1 to 131526)
AUTHORS	Bonning, B. C. and Harrison, R. L.
TITLE	Direct Submission
JOURNAL	Submitted (28-AUG-2002) Entomology, Iowa State University, Ames, Iowa 50011, USA
FEATURES	<p>Location/Qualifiers</p> <p>1..131526</p> <p>/organism="Rachidplusia ou MNPV"</p> <p>/vifion</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:80366"</p> <p>join(131515..131526,1..326)</p> <p>/note="consists of 30 bp imperfect palindromes; hrl; replication origin"</p> <p>392..898</p> <p>/note="ptp; r01; ac1-like"</p>
enhancer	<p>/codon_start=1</p> <p>/product="protein tyrosine/serine phosphatase"</p> <p>/protein_id="AAN28109.1"</p> <p>/db_xref="GI:23476562"</p> <p>/translation="MPPSRKRNHYIQQGVYKDSNLICPKPTLRPELRAYTSSEEDVWTAEQVYKQNPNGALIIDLNTSKSYDGFAPLAGLLYKIQVGGTLPPESIVQEDIDTVKEETKCPGMLVGVHCTGHGINTRYGVWVCYLVMTHTGLIAPQEAIDREKAGKHIERCNYVDLLI"</p>
CDS	<p>complement(962..1417)</p> <p>/note="r02; ac4-like"</p> <p>/codon_start=1</p> <p>/product="unknown"</p> <p>/protein_id="AAN28116.1"</p> <p>/db_xref="GI:23476569"</p> <p>/translation="MKLTYYKASLLKXALSLTPREKENIIPHPHLTRLRLIDGMIRNEDVQRNRNRSRLDISACMQINVHTMPNATIDMRQPNCTIRYQICHLSEADVPSPDNISVRYRLCVACGATPLVIDHPIDVFGHTEEGVNLLEVQRINAGDL"</p> <p>1446..1775</p> <p>/note="r03; ac5-like"</p> <p>/codon_start=1</p> <p>/product="unknown"</p> <p>/protein_id="AAN28124.1"</p> <p>/db_xref="GI:23476577"</p> <p>/translation="MYRTSRINNAPVVASOHYDVRDOIKEBLNSLRNVHDLCTRSGETSFDCKFLRSDMTVPVTTITPPKRTADYKITEYGVDKTIKPSDRPLVESGFLRREAAKYGCIV"</p> <p>1756..2368</p> <p>/note="lef-2; r04; ac6-like"</p> <p>/codon_start=1</p> <p>/product="late expression factor - 2"</p> <p>/protein_id="AAN28095.1"</p> <p>/db_xref="GI:23476548"</p> <p>/translation="MANAKVYVSPLSASGLDKAKYILIDPDPIKTLTPYTVFVNGGLIKISGRLYMLTAPPTINEIKSNFKRSKRNICKCEAKGKAVYDMLNSINMPCCIKILIGLDKENNVPRGMYRRFILTNCYIAVNVSCAKENRKLALTHIFYHDSQGEVEMHLLKSDQVYKCPNCGQMKQNVKDLCPFAGNCKGILNICY"</p> <p>complement(2429..3064)</p> <p>/note="orf603; r05; ac7-like"</p> <p>/codon_start=1</p> <p>/product="unknown"</p> <p>/protein_id="AAN28093.1"</p> <p>/db_xref="GI:23476546"</p> <p>/translation="MAVIFPNKKOLADDSIENGGEFLPLNGSYSLERYVAVPLKKGKVNPDALFKNIDLEPAAYAGNGLIKYDQPSIDVYINLIKATHEEELPENSITVNFTRTMRGTHPIKDKIYIYIDNKKFTLYDYIYIGYDNNYVNFEEKKEKEKEVEDRQIKASSICEKNIISQINCSESFENDFHYLDYIDYAFSIIDNSTVLAFAGLYC"</p> <p>3230..3967</p> <p>/note="polh; r06; ac8-like"</p> <p>/codon_start=1</p> <p>/product="polyhedrin"</p> <p>/protein_id="AAN28081.1"</p> <p>/db_xref="GI:23476534"</p> <p>/translation="MPDYSYRPTIGRTYVYDNNKYKNGSVIRNAKAKKHLIEHEEBE</p>
CDS	



[illegible]

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28161.1"  
 /db\_xref="GI:23476614"  
 /translation="MLFTIIVLVFLRGALCGCALSAFPIAHQPKQVIIVPSKXANGQ  
 NGVCSNALMCO"  
 11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28031.1"  
 /db\_xref="GI:23476523"  
 /translation="MTLCMLALLSTLTVANAAILAVPTPAYSHIIVKYIETALA  
 EKANHVIVPEKLFPEYTKGCGNITENADSDVOQKLVANSAFRRGVSDPTD  
 VTANVYGLITFMFKDPFNINVRNFIANNQDTLVVFAFVYALVFGHLDPAPDIO  
 IAPGGLAENFDYGVAVARHPYHINIRNNDDTBEANVTEMLYKERKILANMNA  
 LNKQGFNPTFTEKLRNKQDLNLTNHPITFNKNRPPVPSYTLGGDHYLVKRASTK  
 LNPINQAQMSKSGTIVYFSSIDTYSFAEPFLYMLINTKTKTIDYITLWIKIDEB  
 VKNTLTPNVTITOMFNQRAVYGHKKNVAFITQGGIODESDELEAQIPMICPMQDQ  
 FYAHKILQQLGVALPDYFVSSDQLTVAINMLNAPAPYKKGMELVALINDKATF  
 PLPDIAIKETERIVRYRHDISRQVLSLKTAAVPSYNYMKVSIVMNLTHP"

CDS

11711..12388  
 /note="rol4; ac16-like"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA28087.1"  
 /db\_xref="GI:23476540"  
 /translation="MESYQTRLCASSNQPAFPRKQQLAVPVSUNSLTHITISTVTS"

Query Match 18.0%; Score 101.8; DB 10; Length 131526;  
 Best Local Similarity 77.0%; Pred. No. 2.2e-23;  
 Matches 124; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

CDS

11578..11745  
 /note="egt; ac15-like"  
 /codon\_start=1  
 /product="ecysteroid UDP-glucosyltransferase"  
 /protein\_id="AA28



```

/note="BRO: 37769 Da primary translation product"
/codon_start=1
/product="baculovirus repeated ORF"
/protein_id="AA06632.1"
/db_xref="GI:559071"
/translation="MARVKIGEFKGFEDFTNLRYLERDOQVRFVAKDVANSKYTYVC
DKAIRVDNKKYSLFEQTLONGPTSNVYKGGDPYLOPHVTLITKSGVLOLIMS
KLAIVAILOEMLEBEVIOYLCGTGKVDPAIKORBEESKOLVTKLIAFTEHTALQAV
VAKTEBELYKQGFIERIVAIKQKQIAKLOQTRVTDLNRMITGQETMOKKDEIM
QKQADQVTDLVAKVLDSDRAVOYPAUKRHPVLCTVRDGTTFATITGQKTYVENOKH
KKNINAVNIIVENIRPNPVTDMNNAIDRLQAKRSKRSIVLVRWKKRNMLKIG"
/gene="Ac-ctx"
/complement (2084 . .2245)
/note="synonym: AcOrf-3"
/complement (2084 . .2245)
/gene="Ac-ctx"
/note="CTX: 5590 Da primary translation product"
/codon_start=1
/product="conotoxin-1like peptide"
/protein_id="AA06633.1"
/db_xref="GI:559072"
/translation="WQIKTVLLAPMFMALNAQHVLAACETGAVCVHDECCSGACS
PIFYCLPQ"
/complement (2295 . .2750)
/gene="AcOrf-4"
/complement (2295 . .2750)
/gene="AcOrf-4"
/note="1757 Da primary translation product"
/codon_start=1
/product="AcOrf-4 peptide"
/protein_id="AA06634.1"
/db_xref="GI:559073"
/translation="MKLTYSASLILKALRLTREYKENIIPHPDHLTRLIDLIDGNIK
SEVQFNRTRNDLISACQIVNRVTMPATIDMKQPCITFRIQQCHLEADVPS
PDHVSYYRLCVACGTPVLDHPLDFGHTBEGVBLEVQRVNASGEL"
/2779 . .3108
/gene="AcOrf-5"
/2779 . .3108
/gene="AcOrf-5"
/note="12435 Da primary translation product"
/codon_start=1
/product="AcOrf-5 peptide"
/protein_id="AA06635.1"
/db_xref="GI:559074"
/translation="MYRTSRINAPVVASQHDYDROIKRELNSLRNVHDLCTRGCT
SPCNKFLNSDDMTPVTTTPKRTADYKITEYGVGVTIKPSNRPLVBSGPLVRBA
KKGECIV"
/3089 . .3721
/gene="Ac-1ef2"
/note="synonym: AcOrf-6"
/3089 . .3721
/gene="Ac-1ef2"
/note="LBF2: 23926 Da primary translation product"
/codon_start=1
/product="late expression factor 2"
/protein_id="AA06636.1"
/db_xref="GI:559075"
/translation="MANASYNWSPILIRASCLODKKATYLLDPDFIDKLTLPYTVFY
NGGVLVKISGLRLMLTAIPTINEIKNSIFKRSRNIQKCEKVEGKRVVDMLNK
INMPICIKLINDIKENNVRGMYRRKRPFLINCYIANVASCACENRCILIKALTHRYN
HBSKCVGEVWHLLIKSDVYKRPNCQKMTVDLCPFAGKCKLANPICNY"
/complement (3759 . .4364)
/gene="Ac-ORF603"
/note="synonym: AcOrf-7"
/complement (3759 . .4364)
/gene="Ac-ORF603"
/note="ORF603 peptide; 23612 Da primary translation
product"
/codon_start=1
/protein_id="AA06637.1"
/db_xref="GI:559076"
/translation="MAVIFNNKQLADNSIEKGELPLFNGSYNILESYYNPVYLKNG
VIELEBAAYVAGNILYKTDPKFIDYINILIKATHSEELPENSTVNVNRKTRMSGTIIH

```

```

PIKDIYIDNKKFTLYDRYIGYDNNVNYFYEKNEKEKEVEEEDDKASLSCENKII
LSQINCESPENDKXYLLSDYNVAFSIIDNTNVVAFGLR"
/gene="Ac-ph"
/note="synonym: AcOrf-8"
/4520 . .5257
/gene="Ac-ph"
/note="polyhedrin; 28642 Da primary translation product"
/codon_start=1
/product="major occlusion body protein"
/protein_id="AA06638.1"
/db_xref="GI:559077"
/translation="MPDYSRPTIGRTYVYDNNKYKLGAVIKAKRKHFAHEIEE
ATIDPLDNTVAEDPFLGPKQKQKLTFLKIRINVKPDTMKLVYGMKGFYRETWRP
MEDSPFIVNDQEVMDVFLVNNKPTRNKCCTFGLAQLARCDVDYVPHVIRIVEPSW
VGSNNERYSLAKKGGCCPIMNLHSEYTSFEOFDRIWENFYKPIVYIGTDSABEE
EILLEVLVFKVEFAPDAPLFTGPAY"
/complement (5287 . .6918)
/gene="Ac-ORF1629"
/note="synonym: AcOrf-9; Ac-61k; Ac-vp78"
/complement (5287 . .6918)
/gene="Ac-ORF1629"
/note="61k; vp78; required for virus replication; 60713 Da
primary translation product"
/codon_start=1
/product="viral capsid associated protein"
/protein_id="AA06639.1"
/db_xref="GI:559078"
/translation="MTNRREYVSQSYLPNNRNKIDAHOFERVDTAEOIIRNITVD
NTVLRNDVLNLIKLANDVPDKAVYVDDSDSVSHNAYVAKRVLGVDRPSLRQ
SLNTYAYIERLNTIGTVNDSERTMLIADPYLNVNVIELPPRALPSRPSVYOP
AAPAVPTTVREBQTKPEQIIPAAPPPSPSPVNIIPAPPPPSKSELPPAPMP7BP
QPAALDDQOQLBEAIRNEKNRTRLPVSKTAPERTTVEVTVLPKTEFEKPSA
SEPPPEPPPPAPAPPPMVDLSAPPPPLVDLPSEMAKPPAPLSNVLSELKSGT
VLKPEQKROSEIIPKSSSTNLADVLADTINRRVAAKSSSASDSEGDGDDN
RPNKANTPVKTYQALFNVTSSOLYTDSEDRNTAKHILNDVEPLDNKQIOTNIDK
ARLLQDILASPVASLSENPIDSPRIGSEKQPLETNRNLFYKSIDIEDLIFRKYDAENH
LIFALTYHKDVKFNEELKTYQQLSNQORTSSA"
/6917 . .7735
/gene="Ac-pk-1"
/note="synonym: AcOrf-10"
/6917 . .7735
/gene="Ac-pk-1"
/note="PK1: 31978 Da primary translation product"
/codon_start=1
/product="protein kinase"
/protein_id="AA06640.1"
/db_xref="GI:559079"
/translation="MATTVATQTLVQVPEKNCNKTRYKINGRGRKISLSHKPTS
KIYQKTIASHNPNADETIKVQDMSDHPFIKTYFNHGSINNQVIMDYIDCDLJET
IQIKGELSYQVLSNIIRQLCEALNDLHKRNFHNDIKLENVYFELADRVYCDYGLC
KHENSLSVHDTGLTEYFSPERIKHTTMHVSFDVYAVGVLYKLTGGRHPFEKSEBML
DLSMKRQOQYNDIGVLKRVNVNARDFYCYLTRVNDRLNVIQIIGHFLS"
/7747 . .7864
/standard_name="h1a"
/note="2 copies of 30 bp imperfect palindromic sequence"
/function="enhancer; replication origin"
/rpt_type="dispersed
complement (7899 . .8921)
/gene="AcOrf-11"
/complement (7899 . .8921)
/gene="AcOrf-11"
/note="40093 Da primary translation product"
/codon_start=1
/product="AcOrf-11 peptide"
/protein_id="AA06641.1"
/db_xref="GI:559080"
/translation="WSLAKKLITVYIYAKNEVHDVYGESHTHYRIYQVRYLSESYNG
KSCIERDVAMRRLKSGSCTPDBAVKMIDAGSISLSHSWSTSTMGIDNVREYLE
QIDAVPVSVQVQNMQIFSLNPFREISQMDLQIILGFEFYFMRKGLRIANV
FNPNDVVGWNTKKEFCVTVYHRIYRISVPALEVRLSAVKVFTRLRKSDVDRLHV
DSYPCRVIAIAYGRFCGIGKEHSKHLSCMHLIFCYLRKCTOEBGSPCYRYIK
DGRQCKDVYKLVKVPDLIAHNSMDKXSLMMLLCVMDCEBIDVDCFYIIFPSFL

```

gene	NMKHLQ"			
	8958..	.9611		
Query Match	18.0%;	Score 101.4;	DB 10;	Length 133894;
Best Local Similarity	75.4%;	Pred. No. 3e-23;		
Matches 126;	Conservative 0;	Mismatches 41;	Indels 0;	Gaps 0;
Qy	112	AACTTTTGGCACTGCAAAAACACGCTTTTGCACGGGGCCCATCATAGTACAAACT 171		
Db	132384	AATTTTTTGCATGCAAAAAGTTACTTTTGCCTGACACTCATATACATTCATCT 132325		
Qy	172	CTACGTTTCTAGACTATTTTACATAATAGCTACACCGTTGTATACGCTCCAAATACA 231		
Db	132324	CTCAAAATCGTAGACTATTTTATAGATAGTACACTAGACATTCAGCTCCCAATATA 132265		
Qy	232	CTACCAACACTGAACTTTTTCAGTGCACAAAAGTACGTCTCGGC 278		
Db	132264	CTACTACACTATCAACTTTTTCATTACAAAAGTTCATTTTTCG 132218		
RESULT 14				
LOCUS	AY327402	131160 bp	DNA	circular VRL 25-MAR-2005
DEFINITION	Choristoneura fumiferana defective nucleopolyhedrovirus complete genome.			
ACCESSION	AY327402	AF068194 U10476 U23422 U78194		
VERSION	AY327402.2	GI:58013199		
KEYWORDS				
SOURCE	Choristoneura fumiferana DEF MNPV (CFDEFMNPV)			
ORGANISM	Choristoneura fumiferana DEF MNPV			
REFERENCE	Virusess; dsDNA virusess, no RNA stage; Baculoviridae;			
AUTHORS	1 (bases 11500 to 13512)			
TITLE	Nucleopolyhedrovirus.			
JOURNAL	Barrett,J.W., Krell,P.J. and Arif,B.M.			
PUBMED	Characterization, sequencing and phylogeny of the ecdysteroicid			
REFERENCE	UDP-glucosyltransferase gene from two distinct nuclear polyhedrosis			
AUTHORS	viruses isolated from Choristoneura fumiferana			
TITLE	J. Gen. Virol. 76 (PT 10), 2447-2456 (1995)			
JOURNAL	7595348			
PUBMED	2 (bases 13243 to 14241)			
REFERENCE	Barrett,J.W., Lauzon,H.A., Mercuri,P.S., Krell,P.J., Sohi,S.S. and			
AUTHORS	Arif,B.M.			
TITLE	The putative LfF-1 proteins from two distinct Choristoneura			
JOURNAL	fumiferana multiple nucleopolyhedroviruses share domain homology to			
PUBMED	eukaryotic primases			
REFERENCE	Virus Genes 13 (3), 229-237 (1996)			
AUTHORS	9035367			
TITLE	3 (bases 80268 to 81693)			
JOURNAL	Li,X., Lauzon,H.A., Sohi,S.S., Palli,S.R., Retnakaran,A. and			
PUBMED	Arif,B.M.			
REFERENCE	Molecular analysis of the p48 gene of Choristoneura fumiferana			
AUTHORS	multicapsid nucleopolyhedroviruses CfMNPV and CfDEFMNPV			
TITLE	J. Gen. Virol. 80 (PT 7), 1833-1840 (1999)			
JOURNAL	10423153			
PUBMED	4 (bases 1 to 131160)			
REFERENCE	Li,X., Barrett,J., Pang,A., Klose,R.J., Krell,P.J. and Arif,B.M.			
AUTHORS	Characterization of an overexpressed spindle protein during a			
TITLE	baculovirus infection			
JOURNAL	Virology 268 (1), 56-67 (2000)			
PUBMED	10683327			
REFERENCE	5 (bases 1 to 131160)			
AUTHORS	Lauzon,H.A., Jamieson,P.B., Krell,P.J. and Arif,B.M.			
TITLE	Gene organization and sequencing of the Choristoneura fumiferana			
JOURNAL	defective nucleopolyhedrovirus genome			
PUBMED	J. Gen. Virol. 86 (PT 4), 945-961 (2005)			
REFERENCE	15784888			
AUTHORS	6 (bases 1 to 131160)			
TITLE	Lauzon,H.A.M., Jamieson,P.B., Krell,P.J. and Arif,B.M.			
JOURNAL	Direct Submission			
PUBMED	Submitted (20-JUN-2003) Molecular Virology, Great Lakes Forestry			
REFERENCE	Centre, 1219 Queen St. E., Sault Ste. Marie, Ontario P6A 2B5,			
AUTHORS	Canada			
TITLE	7 (bases 1 to 131160)			

AUTHORS	Lauzon,H.A.M., Jamieson,P.B., Krell,P.J. and Arif,B.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-JAN-2005) Molecular Virology, Great Lakes Forestry		
COMMENT	Centre, 1219 Queen St. E., Sault Ste. Marie, Ontario P6A 2B5,		
FEATURES	Canada		
REMARK	Sequence update by submitter		
SOURCE	On Jan 21, 2005 this sequence version replaced gi:37499236.		
1..131160	Location/Qualifiers		
/organism="Choristoneura fumiferana DEF MNPV"			
/mol_type="genomic DNA"			
/db_xref="taxon:74660"			
join(129351..131160,1..20)			
/note="ORF 149; 1629 capsid Op2/Ac9"			
/codon_start=1			
/product="viral capsid associated protein"			
/protein_id="AA091783.1"			
/db_xref="GI:37499384"			
/translation="MDROYQSVKSYLNNNAHAIADAPFQIVAGPEAHNLKRLFG			
RTVKLRATLIDLLKLAENIYADTAYMGVNTIESSRHFPATLKRMLLNVODEHPRN			
VLAIVARIENLRSDAVNDVEITVLSGDFEYESVATROYSTDTLPPPLPPI			
PEITQAFVAPSPPPPAQAFVAPSPPPPAQAFVAPSPPPPAQAFVAPSPPPPAQ			
FVAPSPPPPAQAFVAPSPPPPAQAFVAPSPPPPAQAFVAPSPPPPAQAFVAP			
NGVNLNTTLTKPPVPPPAHLSRPNFVGDKYGTNTPPNCTSPQGVNVPPVAP			
PLVMPPPPVPVPPPPDLNLLDAMVSEPRKGTDRGALPDQIMGATLTKAOPVE			
PPSDLRGMDLNGRTGATLTKGTLEDENLGPKRKGEGILGLVNTLTGSRGIDSE			
RSDVATSESTSGFDSADTRADKASSELKHAHLYNFAKDSKLNVIOKVNSELTKI			
LENVGPDLRSPTAENYKANAAGLYLPQHTLTLPKNLADAPABELVAAADAPQVVO			
IEDLRFAGRYDRAAFIQAVDAPEDKLLKFLVNAQJSTRQ"			
complement(1..738)			
/note="ORF 1; ph Op3/Ac8; major occlusion body protein"			
/codon_start=1			
/product="polyhedrin"			
/protein_id="AA091696.1"			
/db_xref="GI:37499297"			
/translation="MENVSRPTIGRTVYDNNKXKGLSVIRAKRKHLLEHDEE			
KHLDPDHVAVEDPFLGPKNOQLTFEINNVDPYTKLIVNNGSKFLBETWRF			
VEDSPFIVNDQEMVDVFLVILNRPTRPNCYKFLQAHARMCDDVPHFVIRVSEY			
VGMNNRYRISLAKKGCGCPINNINSHSEYFESFVNRVWENFYKPIVYIGDSGEE			
ELILEYSLVFKKEPAPDAPLFTGRAY"			
complement(828..1220)			
/note="ORF 2; Op5"			
/codon_start=1			
/product="unknown"			
/protein_id="AA091742.1"			
/db_xref="GI:37499343"			
/translation="MADSTNRILRLFYRSSKTGASLNOADBECLYDEIRPVGAMN			
KQIETSCRKQAEERASALRVELADRNAMOKAAPYCADDGRWSTLSQOULED			
IDRNKEIVRRIYKLQKQODCLNNKAP"			
complement(1247..1858)			
/note="ORF 3; Lef-2 Op6/Ac6"			
/codon_start=1			
/product="late expression factor 2"			
/protein_id="AA091716.1"			
/db_xref="GI:37499317"			
/translation="MGSAMNPAAGVSLQNSEKYLIDNPDEFGVLTALSPCTVFKQGL			
FHEMSGRLRALTLAPKTEPRKAVLHREKRVCLKAGDGVNLAKLNSLRMPICM			
VIMAEINNAAPRGCMYKRPFCYIGNVYSCTKCKGACACIGALLHLYKNDPKCVG			
EVTHLIRKADVDYKSNCAKKTIVNLCFKAQCKGKPICNF"			
complement(1861..2124)			
/note="ORF 4; Op7/Ac5"			
/codon_start=1			
/product="unknown"			
/protein_id="AA091760.1"			
/db_xref="GI:37499361"			
/translation="MNRPMRTANAPVIVSNHDYDREQIRDLNLSIRSVHELCTRST			
TGPDNRLRIETSSNDVITKSPVILIKNSAATVGOSALVCDEV"			
2171..2608			
/note="ORF 5; Op8/Ac4"			
/codon_start=1			
/product="unknown"			
/protein_id="AA091737.1"			
/db_xref="GI:37499338"			

/translation="MASARFIKSLRVTOGKFNIVADITLHATRALIDGKVTHTDV  
 HRFGFLRNLVLEACMAVNOVYVDPGTILRLNPKTIYERVCOKCHAVADVPAPDHIS1  
 ARYLCAACGMVLDIDHPLVDVGTDEGVNELIEVQRINAGGDL"  
 2647. .3657  
 /note="ORF 6; bro Ac2"  
 /codon\_start=1  
 /product="Baculovirus repeated ORF"  
 /protein\_id="AA091674.1"  
 /db\_xref="GI:37499275"  
 /translation="MAQIKIGPKFGEDVETLRVLDIDIVKPVAKVASSLKNTCD  
 KAIRHYDDKTKTSFEOTIOLGSGSTINLVKRGDPLYLQPHITLVITSGVYQLIMSK  
 LPAVLEQEWLSEVLPVCLCTGKINPAIKQOEENKQLINKLVKTSFSDHNTLQTL  
 LQKQELVKKQOEFLERIVATKQIEAKDIQVTRVMDLRNMTNGFQETQKDEIQ  
 OKDAQVTLVAKVIDLSGRAVOYEDERKHPVLCVTDGTTTPIAOCKAYVNHQKIK  
 RNLAAADIVAEITRPNPVDVMMNATHRLAAKSKRTISFSDQAQGFARIRIGLNLN  
 PD"  
 complement (3686. .4168)  
 /note="ORF 7; ptp-2 Op3"  
 /codon\_start=1  
 /product="Protein tyrosine phosphatase 2"  
 /protein\_id="AA091731.1"  
 /db\_xref="GI:37499332"  
 /translation="MTDANOIDENLPIVGYGNNEMALQTIKKHDISSVLSIDSVG  
 PIRQALGPAGYHIVHYCEDEPTCMALINMDALYDIERRINEKKILIHCHAGVSR  
 SATLAVYYMKKQVSEYKALRFVNNKRNVALSDHFVRFLSRCTYAFVNNKLIHVG  
 "  
 complement (4155. .4676)  
 /note="ORF 8; ptp-1 Op10/Ac1"  
 /codon\_start=1  
 /product="Protein tyrosine phosphatase 1"  
 /protein\_id="AA091726.1"  
 /db\_xref="GI:37499327"  
 /translation="MFGRMHNTYTGKVIENGLICFKYPLKQEVPEVYVNDENRMT  
 VSNLVNROPALGAVIDLTNLTIRYDGAKNRDAVYIKKIRPQGEVNPENIEVDFSD  
 VQFSARCPGMLGICVHCTHGLNRTGILVCCYIIDKLHSPDIAAFEARGHKIERR  
 NYEDLLKRVRR"  
 4935. .5921  
 /note="ORF 9; Op11/Ac11"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA091675.1"  
 /db\_xref="GI:37499276"  
 /translation="MSLSQLLVYAYGYEYNNMHERYSESYLRIYINEYLINSYANE  
 TSCVRDIEIARLNGELSFDAAROLNLEIAKLATYHTGEMKSCFSELSVLA  
 EINCABLETGRGASRQVAFALDSLEPPMMDLSQVIMGRFMHPIRGALATHYAVP  
 DEPIKIDGWYNNKFCVLTMYRRIIRGSPRAELTTRIQNTYTKIKVDESSNKLAMG  
 DVGRCGCGKDHFAQHKMSVITFFQYMGVETVADERPPCSVKIDFGROCKETTK  
 DIQLQVILHSHAMTKQKNALFDLLCCNNASDIDVDCYDIYVKKFYNIAYV"  
 6450. .6723  
 /note="4 copies of a 30 bp imperfect palindromic sequence;  
 location similar to Ac h1a"  
 /function="enhancer; replication origin"  
 complement (6723. .7871)  
 /note="ORF 10; Op20/Ac20"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA091663.1"  
 /db\_xref="GI:37499264"  
 /translation="MYRLIVLFLFALLIYIVVPEYQAYOHIQTAQHDYNDLTJDMND  
 YLESVRRHHYVMEALPSVQPTNIGTLAGEFLKCMSPVLYTEIDLPFCTQCE  
 DPQASVFPVGEEDTYVNGHKLAVGYCTNVPDNCNRETSVILMSLQNTIADP  
 RYFAGTSNMTQLAGROHFDRLIPGOSDRNVLFDRLIGREVNATVFNFRSDECDDEA  
 TRFEKRCNARLDNNMLFNPPLNPLECLPNVCTNVAIVYVTSVRPNPFGSCDCCBEA  
 VVRVRYVPGDRSSVCSAYDGDITTTASHRPREVCVNTYTSIGNSNKLLCPSDTF  
 DASTDAAPAEVPGSTPLSGNGIDEPTRHFPDTRIRIXNDVGRGIN"  
 7898. .8863  
 /note="ORF 11; arif Op19/Ac21"  
 /codon\_start=1  
 /product="actin rearrangement inducing factor"  
 /protein\_id="AA091678.1"  
 /db\_xref="GI:37499279"  
 /translation="MLAQINYLILQVLAHALYTTILIAVFSIMGTINRYAFLLELN  
 DNGHSVINLSVLAFLGPCVFTTTMAIYKFLCYKRAMBSNPFYMKTIISIAHMA

Query Match 17.7%; Score 100; DB 10; Length 131160;  
 Best Local Similarity 80.6%; Pred. No. 9.2e-23;  
 Matches 129; Conservative 0; Mismatches 30; Indels 1; Gaps 1;  
 QY 112 AACCTTTTTCACATCGCAAAAAACAGCCTTTTGACAGCGGCCCCATACATG-GTACAAAC 170  
 DB 125539 AACATTATTCAGAGCAAAAAAGTTCACTTTTATTCGACCCCTATATACATACCAAC 125598  
 QY 171 TCTACGTTTGTAGACTATTTTACATTAATAGTCTACACCGTGTATAGCTCTCAATAC 230  
 DB 125599 TCTACGTTTGTAGACTATTTTACATTAATAGTCTACACCGTGTATAGCTCTCAATAC 125658  
 QY 231 ACTACCAACATTTGAACCTTTTGCAGTGCAGAAAAAGTAC 270  
 DB 125659 ACTACCAACATCGAAGCTTTTTCAGCTGCAATTAAGTTTC 125698  
 RESULT 15  
 AY846866/c 1602 bp DNA linear VRL 12-JAN-2005  
 LOCUS Antheraea pernyi nucleopolyhedrovirus occlusion-derived virus  
 DEFINITION Antheraea-56 protein (odv-e56) gene, complete cds.  
 ACCESSION AY846866  
 VERSION AY846866.1 GI:57233514  
 KEYWORDS Antheraea pernyi nucleopolyhedrovirus  
 SOURCE Antheraea pernyi nucleopolyhedrovirus  
 ORGANISM Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 Nucleopolyhedrovirus.  
 REFERENCE 1 (bases 1 to 1602)  
 AUTHORS Xia,A., Li,W., Mu,Z., Li,J., Zhou,Q. and Zhang,Z.  
 TITLE Cloning and analysis of occlusion-derived virus envelope-56 gene of  
 Antheraea pernyi nuclear polyhedrosis virus  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1602)  
 AUTHORS Xia,A., Li,J., Li,W., Zhou,Q. and Zhang,Z.  
 TITLE Direct SubMISSION  
 JOURNAL Submitted (04-DEC-2004) Insect Pathology, The Sarcultural Research  
 Institute Caas, Shabaidu, Zhenjiang, Jiangsu 212018, China  
 FEATURES  
 source location/Qualifiers  
 1..1602  
 /organism="Antheraea pernyi nucleopolyhedrovirus"  
 /mol\_type="genomic DNA"  
 /isolate="Shengyang"  
 /db\_xref="taxon:161494"  
 /country="China"  
 217..1341  
 /gene="odv-e56"  
 217..1341  
 /gene="odv-e56"  
 /note="ODV-E56"  
 /codon\_start=1  
 /product="occlusion-derived virus envelope-56 protein"  
 /protein\_id="AA048304.1"  
 /db\_xref="GI:57233515"  
 /translation="MSFFNLRKRVKVPYSGADPFVSDNLRVYLPSTPAGFTNVLSTPS  
 TNLGGREPPVYNLNNQFVSAAGMNRITRNNDVPRISTPGISDDPQIGSLAOLRR  
 VUNVDPAGHYVKTGDAVKQFPEPTNRSADGVPRALQONRLATYTDGAKTAGV  
 LTAGGYLAFAFSAITVQDITLNLTNGTSYVYRSGSGSTAAACLLSTCRDRPWN  
 TSEVAICADPLVNAVVOLOQICNENFTGIEQTVROSPDAPDPSPQYVDISLVP  
 QTIMCIPEPNLIDGLIDGLIDLLGEDGLIGKSSSDSVSKMLPLIWLIGAVFLG  
 LIIFYRVRVVGSGSGSATATAQPSPIVWQPPAPARE"  
 ORIGIN  
 Query Match 17.4%; Score 98.4; DB 10; Length 1602;  
 Best Local Similarity 82.5%; Pred. No. 2.5e-22;  
 Matches 137; Conservative 0; Mismatches 26; Indels 3; Gaps 2;  
 QY 110 CCAACTTTTTCGACGCAAAAAACAGCCTTTTGACAGCGG-GCCCATACATAGTACAA 168

```
Db      176  CAACTTTTGTGTAAGTCAAAAAACAGTTTGGCGAGGCCCCCATATACAGTACA 117
Qy      169  ACTCTAGCTTTCGTAGACTATTTTACATTAATAGCTACACCGTTGTATAGCTCCAAT 228
Db      116  AGCTACGTTTCGTAGACTATTTTACCTAATAATAGCTACACTGTCTATAGCTGCCAAT 57
Qy      229  ACACTACC--ACACATTGAACCTTTTGCAGTGCAAAAAGTAGCT 272
Db      56  ACACTACCCCTTAACCAAACTATTTTGCAGTGCAAAAAGTAGCT 11
```

Search completed: June 15, 2006, 19:57:54  
Job time : 4072.26 secs



CC virus (OPMNPV) immediate early 2 (Op i2) promoter sequence. The  
 CC invention provides a new shuttle vector for transforming insect cells  
 CC that comprises: (i) prokaryotic origin of replication; (ii) insect  
 CC promoter having homology to, and capable of functioning as, an immediate  
 CC early baculovirus promoter; (iii) prokaryotic promoter sequence, and (iv)  
 CC selectable marker capable of conferring resistance to a bleomycin/  
 CC phleomycin-type antibiotic under transcriptional control of (ii) and  
 CC (iii), in insect and prokaryotic cells respectively. The vectors can be  
 CC used to stably transform (especially insect) cells with heterologous DNA,  
 CC useful to allow study of gene expression and direct expression of  
 CC heterologous gene products e.g. commercially important proteins. They are  
 CC especially useful to allow expression of the heterologous  
 CC melanotransferrins, ion transport peptide hormones or biologically active  
 CC derivatives in insect cells. They enable transformation of insect cell  
 CC lines from disparate species, allowing screening of lines for optimum  
 CC post-translational modification of particular proteins. Shuttle vectors  
 CC further comprising DNA transposable elements defining a transposon can be  
 CC used to optimise heterologous protein expression and facilitate selection  
 CC of desired transformants. (Updated on 17-OCT-2003 to standardise OS  
 CC field)

SQ Sequence 564 BP, 147 A, 160 C, 116 G, 141 T, 0 U, 0 Other;

Query Match Best Local Similarity 100.0%; Score 564; DB 2; Length 564;

Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATGATGCTAATGTTGTTCAACAACAATCTGTGAACGTG 60  
 DB 1 CATGATGATTAACAATGATGCTAATGTTGTTCAACAACAATCTGTGAACGTG 60  
 QY TTTTCATGTTGGCCAAAGACCTTATATCTCGGTGGCCCTCCCAACAACCTTTT 120  
 DB TTTTCATGTTGGCCAAAGACCTTATATCTCGGTGGCCCTCCCAACAACCTTTT 120  
 QY 121 GCACTGCAAAAAAACAAGCTTTTGCACGCGGCGCCATACATAGTACAACTGATT 180  
 DB 121 GCACTGCAAAAAAACAAGCTTTTGCACGCGGCGCCATACATAGTACAACTGATT 180  
 QY 181 GTGACCTATTTTACATAAATAGTCTACACCGTTTATACGTCCTCAATACATAC 240  
 DB 181 GTGACCTATTTTACATAAATAGTCTACACCGTTTATACGTCCTCAATACATAC 240  
 QY 241 ATTGAACCTTTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 300  
 DB 241 ATTGAACCTTTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 300  
 QY 301 GGGTCGCGTCTGTACGATGCAATCATTTATCGAAGCGAGGAGTGTCTTATCGT 360  
 DB 301 GGGTCGCGTCTGTACGATGCAATCATTTATCGAAGCGAGGAGTGTCTTATCGT 360  
 QY 361 GACAGAGCGGCAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
 DB 361 GACAGAGCGGCAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
 QY 421 GGAAGCGGCTCCATATACGCGCGCTTATCTATGCGCGTACCGAGCGAGCGCC 480  
 DB 421 GGAAGCGGCTCCATATACGCGCGCTTATCTATGCGCGTACCGAGCGAGCGCC 480  
 QY 481 GTCCCGCTTATCGGCTTATTAATACAGCCGCAACATCTGTAAACAGTTGAACAG 540  
 DB 481 GTCCCGCTTATCGGCTTATTAATACAGCCGCAACATCTGTAAACAGTTGAACAG 540  
 QY 541 CATCTGTTACAGGACCAACATG 564  
 DB 541 CATCTGTTACAGGACCAACATG 564

RESULT 2  
 ADQ48575  
 ID ADQ48575 standard; DNA, 560 BP.  
 XX  
 AC ADQ48575;

XX 09-SEP-2004 (first entry)  
 DT  
 XX  
 DE OpIE2 promoter DNA sequence.  
 XX  
 XX viral vector; recombinant site; recombinant virus;  
 KW replication-defective particle generation; gene expression inhibition;  
 KM gene therapy vector; de; OpIE2 promoter.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004009768-A2.  
 XX  
 PD 29-JAN-2004.  
 XX  
 PF 18-JUL-2003; 2003WO-US022437.  
 XX  
 PR 18-JUL-2002; 2002US-0396335P.  
 PR 26-JUL-2002; 2002US-0398617P.  
 PR 19-NOV-2002; 2002US-0427231P.  
 PR 24-MAR-2003; 2003US-0456496P.  
 PR 03-JUN-2003; 2003US-0474940P.  
 XX  
 PA (INVI-) INVITROGEN CORP.  
 PA (BENN/) BENNETT R P.  
 PA (WELC/) WELCH P J.  
 PA (HARM/) HARWOOD S.  
 PA (MADD/) MADDEN K.  
 PA (FRIM/) FRIMPONG K.  
 PA (FRAN/) FRANKIE K E.  
 XX  
 PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;  
 XX WPI; 2004-132944/13.  
 XX  
 XX Novel nucleic acid molecule comprising all or a portion of one or more  
 PT viral genome and further comprising two or more recombination sites that  
 PT do not substantially recombine with each other, useful as gene therapy.  
 XX  
 PS Disclosure; Fig 16; 555bp; English.  
 XX  
 XX The invention comprises a nucleic acid molecule consisting of all or a  
 CC portion of one or more viral/baculoviral genomes and further containing  
 CC at least two recombination sites that do not substantially recombine with  
 CC each other. One or more of the recombination sites is capable of  
 CC undergoing recombination with a compatible recombination site in the  
 CC presence of one or more proteins active in lambda recombination. The  
 CC nucleic acid of the invention replicates in prokaryotic and eukaryotic  
 CC cells. The nucleic acid of the invention is useful for constructing a  
 CC recombinant virus, generating replication-defective particles,  
 CC preventing/inhibiting the expression of one or more genes in an organism,  
 CC and are useful as gene therapy vectors. The nucleic acid of the invention  
 CC is also useful for producing and expressing fusion polypeptides. The  
 CC present DNA sequence represents the OpIE2 promoter that was used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 560 BP, 144 A, 156 C, 116 G, 144 T, 0 U, 0 Other;

Query Match Best Local Similarity 97.2%; Score 548; DB 12; Length 560;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATGATGCTAATGTTGTTCAACAACAATCTGTGAACGTG 60  
 DB 5 CATGATGATTAACAATGATGCTAATGTTGTTCAACAACAATCTGTGAACGTG 64  
 QY TTTTCATGTTGGCCAAAGACCTTATATCTCGGTGGCCCTCCCAACAACCTTTT 120  
 DB TTTTCATGTTGGCCAAAGACCTTATATCTCGGTGGCCCTCCCAACAACCTTTT 124  
 QY 121 GCACTGCAAAAAAACAAGCTTTTGCACGCGGCGCCATACATAGTACAACTGATT 180  
 DB 125 GCACTGCAAAAAAACAAGCTTTTGCACGCGGCGCCATACATAGTACAACTGATT 184



OY	181	GTAGACATATTTACATPAAATAGCTCAACCGTGTATAGCTCCCAATACACTACACAC	240
Db	185	GTAGACTATTTACATPAAATAGCTCAACCGTGTATAGCTCCCAATACACTACACAC	244
OY	241	ATTGAACCTTTTGTGAGTGCACAAAAGTACGTGTGCGACGTACGTAGGCGGCTTATC	300
Db	245	ATTGAACCTTTTGTGAGTGCACAAAAGTACGTGTGCGACGTACGTAGGCGGCTTATC	304
OY	301	GGGTGGCGTCTCTGTCACTACGATCACTTATCGAACCGGACGAGTGTGTCTTATCGT	360
Db	305	GGGTGGCGTCTCTGTCACTACGATCACTTATCGAACCGGACGAGTGTGTCTTATCGT	364
OY	361	GACGAGACGCGACGCTTCTGTGTGTCTAACCGGAGCGGACGCACTCTTATCGGAACA	420
Db	365	GACGAGACGCGACGCTTCTGTGTGTCTAACCGGAGCGGACGCACTCTTATCGGAACA	424
OY	421	GGACGCGCTCTCATATACAGCCGCGCTTATCTCATGCGCGGTGACCGGACACGAGCGCCC	480
Db	425	GGACGCGCTCTCATATACAGCCGCGCTTATCTCATGCGCGGTGACCGGACACGAGCGCCC	484
OY	481	GTCCCGCTTATGCGGCTTATPAAATACGCCCGCAACGATCTGGTAAACACAGTTGAACAG	540
Db	485	GTCCCGCTTATGCGGCTTATPAAATACGCCCGCAACGATCTGGTAAACACAGTTGAACAG	544
OY	541	CATCTGTT 548	
Db	545	CATCTGTT 552	

	RESULT 3
AAL61306	AAL61306 standard; DNA; 2773 BP.
ID AAL61306	
XX AAL61306;	
AC	
XX	
D7	22-SEP-2003 (first entry)
XX	
DE pZOP2F expression vector for insect cells.	
XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;	
KW tumour necrosis factor alpha; gene therapy; arthritis; gene; de.	
XX	
OS unidentified.	
XX WO2003042244-A2.	
PN	
XX	
PD 22-MAY-2003.	
XX	
PF 15-NOV-2002; 2002WO-DK000764.	
XX	
PR 16-NOV-2001; 2001DK-000011702.	
FR 16-NOV-2001; 2001US-0331575P.	
XX	
PA (PHAR-) PHARMEXA AS.	
PA (KLVS/) KLYSNER S.	
PA (NIEL/) NIELSEN F S.	
PA (BRAT/) BRATT T.	
PA (VOLD/) VOLDBORG B.	
PA (MOUR/) MOURITSEN S.	
P1 Klynsner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;	
PI WPI; 2003-449558/42.	
DR	
XX	
PT New immunogenic analogue of a polymeric protein, useful for preparing a	
XX composition for treating inflammatory diseases e.g. arthritis.	
PS Disclosure; Page 195-196; 196pp; English.	
XX	
CC The invention relates to immunogenic analogues of multimeric proteins	
CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis	
CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic	

Sequence 2773 BP, 669 A, 750 C, 673 G, 681 T, 0 U, 0 Other;  
 CC analogues. The immunogenic analogue is useful for preparing a composition  
 CC for treating inflammatory diseases, e.g., arthritis. It is also used in  
 CC gene therapy. The present sequence is p220p2F expression vector for  
 CC insect cells. This sequence is used to illustrate the method of the  
 CC invention.

Query Match	97.2%;	Score 548;	DB 9;	Length 2773;
Best Local Similarity	100.0%;	Pred. No. 4.4e-167;		
Matches 548; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	CATGATGATAAACAATGATAGGTGCTAATGTTGCTTCAACAACAATCTGTGAACTGTG	60
Db	5	CATGATGATAAACAATGATAGGTGCTAATGTTGCTTCAACAACAATCTGTGAACTGTG	64
Qy	61	TTTTCAATGTTGGCAACAAGACCTTTATAGTCGGTGGCTCCCGACACAACTTTTTT	120
Db	65	TTTTCAATGTTGGCAACAAGACCTTTATAGTCGGTGGCTCCCGACACAACTTTTTT	124
Qy	121	GCACATGCAAAAAACACGCTTTTGGACGGGGGCCATACATAGTACAACTTACGTTTC	180
Db	125	GCACATGCAAAAAACACGCTTTTGGACGGGGGCCATACATAGTACAACTTACGTTTC	184
Qy	181	GTAGCATATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACACAC	240
Db	185	GTAGCATATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACACAC	244
Qy	241	ATTGAACCTTTTGGCAGTGCAAAAAAGTACGTGTGCGCAGTCACTGAGCGCGCTTATC	300
Db	245	ATTGAACCTTTTGGCAGTGCAAAAAAGTACGTGTGCGCAGTCACTGAGCGCGCTTATC	304
Qy	301	GGGTGCGGTCTGTCACTGACGAATCACATTATGGACCGGACGAGTGTGTCTTATCGT	360
Db	305	GGGTGCGGTCTGTCACTGACGAATCACATTATGGACCGGACGAGTGTGTCTTATCGT	364
Qy	361	GACAGGAGCGCAGCTTCCGTGTGTGCTTACCGGACCGGACCGCAACCTCTTATCGGAACA	420
Db	365	GACAGGAGCGCAGCTTCCGTGTGTGCTTACCGGACCGGACCGCAACCTCTTATCGGAACA	424
Qy	421	GGACGGCGCTCATATTCAGCGCGGCTTATCTCATGTGCGGTGACCGGACACGAGGGCGCC	480
Db	425	GGACGGCGCTCATATTCAGCGCGGCTTATCTCATGTGCGGTGACCGGACACGAGGGCGCC	484
Qy	481	GTTCGGCTTATCGGGCTATATAATACAGCCGCAACGATCTGTAAACAAGTTGAACAG	540
Db	485	GTTCGGCTTATCGGGCTATATAATACAGCCGCAACGATCTGTAAACAAGTTGAACAG	544
Qy	541	CATTCGTT	548
Db	545	CATTCGTT	552

	RESULT 4
ADQ48539	
ID	ADQ48539 standard; DNA; 5038 BP.
XX	
AC	ADQ48539;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	viral vector-related plasmid - p1B/V5-His-DEST.
XX	
KM	viral vector; recombination site; recombinant virus;
KM	replication-defective particle generation; gene expression inhibition.
KM	gene therapy vector; ds, plasmid.
XX	
OS	unidentified.
XX	
PN	WO2004009768-A2.
XX	
PD	29-JAN-2004.
XX	

PF 18-JUL-2003; 2003WC-US022437.  
 XX 18-JUL-2002; 2002JUS-0396335P.  
 PR 26-JUL-2002; 2002JUS-0398617P.  
 PR 19-NOV-2002; 2002JUS-0427231P.  
 PR 24-MAR-2003; 2003JUS-0456496P.  
 PR 03-JUN-2003; 2003JUS-0474940P.  
 XX (INV-) INVITROGEN CORP.  
 PA (BENN/) BENNETT R P.  
 PA (WELC/) WELCH P J.  
 PA (HARW/) HARWOOD S.  
 PA (MADD/) MADDEN K.  
 PA (FRIM/) FRIMPONG K.  
 PA (FRAN/) FRANKE K E.  
 PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;  
 XX WPI; 2004-132944/13.  
 XX  
 PT Novel nucleic acid molecule comprising all or a portion of one or more  
 PT viral genome and further comprising two or more recombination sites that  
 PT do not substantially recombine with each other, useful as gene therapy.  
 XX  
 PS Example 18; Page 395-403; 555pp; English.  
 XX  
 CC The invention comprises a nucleic acid molecule consisting of all or a  
 CC portion of one or more viral/baculoviral genomes and further containing  
 CC at least two recombination sites that do not substantially recombine with  
 CC each other. One or more of the recombination sites is capable of  
 CC undergoing recombination with a compatible recombination site in the  
 CC presence of one or more proteins active in lambda recombination. The  
 CC nucleic acid of the invention replicates in prokaryotic and eukaryotic  
 CC cells. The nucleic acid of the invention is useful for constructing a  
 CC recombinant virus, generating replication-defective particles,  
 CC preventing/inhibiting the expression of one or more genes in an organism,  
 CC and are useful as gene therapy vectors. The nucleic acid of the invention  
 CC is also useful for producing and expressing fusion polypeptides. The  
 CC present DNA sequence represents a plasmid that was used in the  
 CC exemplification of the invention.  
 CC  
 SQ Sequence 5038 BP; 1342 A; 1178 C; 1198 G; 1320 T; 0 U; 0 Other;  
 Query Match 96.9%; Score 546.4; DB 12; Length 5038;  
 Best Local Similarity 99.8%; Pred. No. 1.9e-16;  
 Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CATGATGATTAACATATGATGCTAATGTTGCTTCAACAACAATTCGTGAACTGTG 60  
 DB 1 CATGATGATTAACATATGATGCTAATGTTGCTTCAACAACAATTCGTGAACTGTG 60  
 QY 61 TTTTCATGTTTGGCCAAACAGCCTTTATATCTCGGTGGCTCCGCCACCACTTTT 120  
 DB 61 TTTTCATGTTTGGCCAAACAGCCTTTATATCTCGGTGGCTCCGCCACCACTTTT 120  
 QY 121 GCATGCAAAAAAACAAGCGTTTGGACGCGGGCCCATACATATGATGAACTTACGTTTC 180  
 DB 121 GCATGCAAAAAAACAAGCGTTTGGACGCGGGCCCATACATATGATGAACTTACGTTTC 180  
 QY 181 GTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCCAAATACATACACAC 240  
 DB 181 GTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCCAAATACATACACAC 240  
 QY 241 ATTGAACCTTTTGGACGTGCAAAAAAATAGTGTCCGACATCAGTACGCGGCTTATC 300  
 DB 241 ATTGAACCTTTTGGACGTGCAAAAAAATAGTGTCCGACATCAGTACGCGGCTTATC 300  
 QY 301 GGGTCGCGTCTGTGACGATGCAATCATTTATGAGACCGGACGAGTGTCTTATCGT 360  
 DB 301 GGGTCGCGTCTGTGACGATGCAATCATTTATGAGACCGGACGAGTGTCTTATCGT 360  
 QY 361 GACAGGACGCGACGCTTCTGTGTGCTTAAACCGGACGCGGACGCAACTCTTATCGAACA 420

DB 361 GACAGGACGCGACGCTTCTGTGTGCTTAAACCGGACGCGGACGCAACTCTTATCGAACA 420  
 QY 421 GAGCGGCTTCATATCAGCGCGCGCTTATCTATGCGGTACCGGACACGAGGGCCCC 480  
 DB 421 GAGCGGCTTCATATCAGCGCGCGCTTATCTATGCGGTACCGGACACGAGGGCCCC 480  
 QY 481 GTCCCGCTTATCGCGCTTAAATATACAGCCCGCAAGATCTGCTAAACAGTTGAACAG 540  
 DB 481 GTCCCGCTTATCGCGCTTAAATATACAGCCCGCAAGATCTGCTAAACAGTTGAACAG 540  
 QY 541 CATCTGTT 548  
 DB 541 CATCTGTT 548  
 RESULT 5  
 AAT13730/c  
 ID AAT13730 standard; DNA; 279 BP.  
 XX  
 AC AAT13730;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 05-SEP-1996 (first entry)  
 XX  
 DE ACNPV ORF 152, residues 132387-132109.  
 XX  
 KW Autographa californica nuclear polyhedrosis virus clone 6; disruption;  
 KW non-essential gene; heterologous protein production; expression vector;  
 KW baculovirus; ss.  
 XX  
 OS Autographa californica nucleopolyhedrovirus; clone 6.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..279  
 FT /tag= a  
 FT /number= ORF 152  
 FT /note= "corresponds to ACNPV nucleotides complement  
 FT (132387-132109)"  
 XX  
 PN W09601320-A2.  
 XX  
 PD 18-JAN-1996.  
 XX  
 PF 30-JUN-1995; 95MO-IB000578.  
 XX  
 PR 04-JUL-1994; 94GB-00013420.  
 XX  
 PA (NATU-) NATURAL ENVIRONMENT RES COUNCIL.  
 PI Bishop D, Possee R, Ayres M;  
 XX WPI; 1996-087670/09.  
 DR  
 XX Autographa californica nuclear polyhedrosis virus complete genome  
 PT sequence - useful in the prodn. of vectors for enhanced heterologous  
 PT protein expression, such as interleukin(s), interferon(s) and  
 PT neurotoxin(s).  
 PS Claim 1; Page 90-186; 122pp; English.  
 XX  
 CC AAT13636-731 show open reading frames 13, 20, 22-30, 32, 38, 41-46, 50-  
 CC 62-63, 66, 68-79, 81-87, 91-92, 96-98, 101-103, 106-127, 129-130, 140  
 CC - 146, 148-150, 152 and 154 from a total of 154 ORFs identified in the  
 CC Autographa californica nuclear polyhedrosis virus (ACNPV) clone 6. Each  
 CC gene is numbered according to its position in the virus genome beginning  
 CC at the left end of the linear map, and irrespective of its orientation.  
 CC The direction of transcription is relative to that of the polyhedrin  
 CC gene. Of the 154 ORFs identified it was found that some of the ORFs (ORFs  
 CC 27, 30, 32, 71, 86, 123, 126 and 127) are dispensable for virus  
 CC replication in cell culture or insect larvae. These genes can be deleted  
 CC from the genome to: (a) provide additional sites for inserting single or  
 CC multiple copies of foreign genes; and (b) to reduce the size of the virus  
 CC complementary strand relative to the polyhedrin gene. The present

CC sequence is designated ORF 152, and is on the complementary strand  
 CC relative to the polyhedrin gene. (Updated on 16-OCT-2003 to standardise  
 CC OS field)

XX Sequence 279 BP; 88 A; 36 C; 68 G; 87 T; 0 U; 0 Other;

Query Match 18.0%; Score 101.4; DB 2; Length 279;  
 Best Local Similarity 75.4%; Pred. No. 2.9e-22;  
 Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 112 AACTTTTGGACGTGCAAAAACACGCTTTTGGACGCGGGCCCATATGATCAAACT 171  
 DB AATTTTTCGATGCAAAAAGTTCACTTTGGCTGACACTCATATAGTACAAATCT 217  
 QY 172 CTACGTTTGGTACATATTTTACATTAATAGCTTACACCGTTGATACGCTCCAAATACA 231  
 DB CTACAAATCGTACATATTTTATATAGATGTCTACACTGTACATGCGCTCCCAATATA 157

QY 232 CTACACACATGTGAACCTTTTGGACGTGCAAAAAGTACGTGCGC 278  
 DB CTACTACACTATCAACTTTTGGCATTTACAAAAGTTCAATTTTGC 110

RESULT 6  
 AAT13635/c  
 ID AAT13635 standard; DNA; 133894 BP.  
 XX  
 AC AAT13635;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 03-SEP-1996 (first entry)  
 XX  
 DE ACNPV genomic DNA clone 6.  
 XX  
 KW Autographa californica nuclear polyhedrosis virus clone 6; disruption;  
 KW non-essential gene; heterologous protein production; expression vector;  
 KW baculovirus; ss.  
 XX  
 OS Autographa californica nucleopolyhedrovirus; clone 6.  
 XX  
 PN WO9601320-A2.  
 XX  
 PD 18-JAN-1996.  
 XX  
 PF 30-JUN-1995; 95MO-IB000578.  
 XX  
 PR 04-JUL-1994; 94GB-00013420.  
 XX  
 PA (NATU-) NATURAL ENVIRONMENT RES COUNCIL.  
 XX  
 PI Bishop D, Possee R, Ayres M;  
 XX  
 DR WPI; 1996-087670/09.  
 XX  
 PT Autographa californica nuclear polyhedrosis virus complete genome  
 PT sequence - useful in the prodn. of vectors for enhanced heterologous  
 PT protein expression, such as interleukin(s), interferon(s) and  
 PT neurotoxin(s).  
 XX  
 PS Disclosure; Page 90-186; 122pp; English.  
 XX  
 CC The complete nucleotide sequence of the genome of clone 6 of the  
 CC baculovirus Autographa californica nuclear polyhedrosis virus (ACNPV) has  
 CC been determined. The sequence is taken from the Genbank record L22858.  
 CC The patent specification claims a polynucleotide selected from open  
 CC reading frames (ORFs 13, 20, 22-26, 28-30, 32, 38, 41-46, 50-60, 62-63,  
 CC 66, 68-79, 81-87, 91-92, 96-98, 101-103, 106-126, 129-130, 140-146, 148-  
 CC 150, 152 and 154 from a total of 154 ORFs identified by the patentees.  
 CC See T13635-731. Expression vectors conrg. the complete genomic sequence  
 CC of ACNPV, with the exception that at least one non-essential ORF is  
 CC disrupted or replaced are useful for the synthesis of heterologous  
 CC proteins. (Updated on 16-OCT-2003 to standardise OS field)

SQ Sequence 133894 BP; 39195 A; 27151 C; 27347 G; 40201 T; 0 U; 0 Other;

Query Match 18.0%; Score 101.4; DB 2; Length 133894;  
 Best Local Similarity 75.4%; Pred. No. 5.2e-21;  
 Matches 126; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 112 AACTTTTGGACGTGCAAAAACACGCTTTTGGACGCGGGCCCATATGATCAAACT 171  
 DB 132384 AATTTTTCGATGCAAAAAGTTCACTTTGGCTGACACTCATATAGTACAAATCT 132325

QY 172 CTACGTTTGGTACATATTTTACATTAATAGCTTACACCGTTGATACGCTCCAAATACA 231  
 DB 132324 CTACAAATCGTACATATTTTATATAGATGTCTACACTGTACATGCGCTCCCAATATA 132265

QY 232 CTACACACATGTGAACCTTTTGGACGTGCAAAAAGTACGTGCGC 278  
 DB 132264 CTACTACACTATCAACTTTTGGCATTTACAAAAGTTCAATTTTGC 132218

RESULT 7  
 ADCS1646/c  
 ID ADCS1646 standard; DNA; 28413 BP.  
 XX  
 AC ADCS1646;  
 XX  
 DT 18-DEC-2003 (first entry)  
 DT  
 XX  
 DE BmNPV genomic DNA nucleotides 100001-128413.  
 DE  
 KW de; silkworm; silkworm movement supression.  
 KW  
 XX  
 OS Bombyx mori nuclear polyhedrosis virus.  
 XX  
 PN JP2003024062-A.  
 XX  
 PD 28-JAN-2003.  
 XX  
 PF 10-JUL-2001; 2001JP-00209305.  
 XX  
 PR 10-JUL-2001; 2001JP-00209305.  
 XX  
 PA (RIKA ) RIKAGAKU KENKYUSHO.  
 XX  
 DR WPI; 2003-516415/49.  
 XX  
 XX  
 XX Novel Bombyx mori polyhedrosis virus in which open reading frame gene is  
 PT inactivated, or its variant, useful for suppressing movement of Bombyx  
 PT mori.  
 XX  
 PS Claim 5; SEQ ID NO 3; 53bp; Japanese.  
 XX  
 CC The invention relates to a Bombyx mori (silk worm) polyhedrosis virus in  
 CC which open reading frame 8 (ORF 8) gene is inactivated. The inactivated  
 CC orf8 gene is useful for suppressing movement of silkworm and efficiently  
 CC prevents the movement of a silkworm from a chamber in which it is raised.  
 CC The present sequence is used in the exemplification of the invention.  
 XX  
 SQ Sequence 28413 BP; 8394 A; 5713 C; 5940 G; 8366 T; 0 U; 0 Other;

Query Match 16.6%; Score 93.4; DB 10; Length 28413;  
 Best Local Similarity 72.5%; Pred. No. 1e-18;  
 Matches 121; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 112 AACTTTTGGACGTGCAAAAACACGCTTTTGGACGCGGGCCCATATGATCAAACT 171  
 DB 22231 AATTTTTCGATGCAAAAAGTTGCGCTATGTTGACATATATATACAGTACGACT 22172

QY 172 CTACGTTTGGTACATATTTTACATTAATAGCTTACACCGTTGATACGCTCCAAATACA 231  
 DB 22171 CTACAAATCGTACATATTTTATATAGATGTCTACACTGTACTAATACGCTTCATATATA 22112

QY 232 CTACACACATGTGAACCTTTTGGACGTGCAAAAAGTACGTGCGC 278

DB 22111 CTACTACATCATCTTTTGGATTACAAAAAGTTCATTTTGCG 22065

# RESULT 8

ADQ48576  
ID ADQ48576 standard; DNA; 141 BP.

AC ADQ48576;

DT 09-SEP-2004 (first entry)

DE Viral vector-related plasmid pIB/V5-His-DEST recombination region #1.

XX viral vector; recombination site; recombinant virus;

KM replication-defective particle generation; gene expression inhibition;

KW gene therapy vector; ds; plasmid; recombination region.

OS Unidentified.

PN WO2004009768-A2.

PD 29-JAN-2004.

PF 18-JUL-2003; 2003WO-US022437.

PR 18-JUL-2002; 2002US-0396335P.

PR 26-JUL-2002; 2002US-0398617P.

PR 19-NOV-2002; 2002US-0427231P.

PR 24-MAR-2003; 2003US-0456496P.

PR 03-JUN-2003; 2003US-0474940P.

XX (INVI-) INVITROGEN CORP.

PA (BENN/) BENNETT R. P.

PA (WELC/) WELCH P. J.

PA (HARM/) HARWOOD S.

PA (MADD/) MADDEN K.

PA (FRIM/) FRIMPONG K.

PA (FRAN/) FRANKS K. E.

XX

PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;

XX WPI; 2004-132944/13.

XX

PT Novel nucleic acid molecule comprising all or a portion of one or more

PT viral genome and further comprising two or more recombination sites that

PT do not substantially recombine with each other, useful as gene therapy.

XX

PS Disclosure; Fig 17; 555pp; English.

XX

XX

XX

XX

Sequence 141 BP; 43 A; 36 C; 27 G; 34 T; 0 U; 1 Other;

Query Match 11.0%; Score 62; DB 12; Length 141;

Best Local Similarity 100.0%; Pred. No. 1.4e-09;

Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

487 CTTATCGCGCCTTAATATACAGCCCGACAGATCTGTTAAACACAGTTGAACAGCATCTG 546

1 CTTATCGCGCCTTAATATACAGCCCGACAGATCTGTTAAACACAGTTGAACAGCATCTG 60

QY 547 TT 548

DB 61 TT 62

# RESULT 9

ABZ10059/c  
ID ABZ10059 standard; DNA; 6289 BP.

AC ABZ10059;

DT 16-JAN-2003 (first entry)

DE Hematopoietic cell proliferation disorder related DNA sequence #199.

XX Human; haematopoietic cell proliferation disorder; cytostatic;

KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;

KW cytosine methylation state; gene; ds.

OS Homo sapiens.

PN WO200277272-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002WO-EP003401.

PR 26-MAR-2001; 2001US-0278333P.

PR (EPIC-) EPIGENOMICS AG.

XX Berlin K, Braun A, Disler J, Guetig D, Howe A, Mueller J;

PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Iesche R, Leu E;

PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Peter C;

PI Schwope I, Ziebarth H;

DR WPI; 2003-018942/01.

XX

PT Detecting and differentiating between hematopoietic cell proliferative

PT disorders, comprises contacting a target nucleic acid with a reagent that

PT distinguishes between methylated and non-methylated CpG dinucleotides.

XX

PS Claim 28; SEQ ID NO 199; 117pp; English.

XX

CC The present invention describes a method for detecting and

CC differentiating between hematopoietic cell proliferative disorders

CC associated with at least 1 gene and/or their regulatory regions in a

CC subject. The method comprises contacting a target nucleic acid in a

CC biological sample obtained from the subject with at least 1 reagent,

CC which distinguishes between methylated and non-methylated CpG

CC dinucleotides within the target nucleic acid. ABZ0961 to ABZ1118

CC represent specifically claimed nucleotide sequences from the present

CC invention. Oligonucleotides from the present invention can be used: for

CC differentiating between healthy haematopoietic cells and proliferative

CC disorder haematopoietic cells; for differentiating between acute

CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for

CC determining the cytosine methylation state and/or single nucleotide

CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder

CC related sequences and their complements; and as primers for the

CC amplification of haematopoietic cell proliferation disorder related DNA

CC sequences. The nucleotide sequences from the present invention can also

CC be used for detecting a predisposition to, differentiation between

CC subclasses, diagnosis, prognosis, treatment and/or monitoring of

CC haematopoietic cell proliferative disorders. The present method enables a

CC highly specific classification of haematopoietic cell proliferative

CC disorders allowing for improved and informed treatment of patients

XX

Sequence 6289 BP; 1680 A; 178 C; 1488 G; 2943 T; 0 U; 0 Other;

Query Match 7.0%; Score 39.6; DB 8; Length 6289;

Best Local Similarity 53.2%; Pred. No. 0.16;

Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTGGACGCAAAAACACGCTTTTGACGCGGGCCATACATAGTCAAACT 171  
 DB 4896 AATTATTTATATTAACCAAAATTAACCTATTAATCACTACCAAAATCATTCATAAAA 4837  
 QY 172 CTACGTTTCGTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCAAAATCA 231  
 DB 4836 TAATTAATTTTAAACATTAATTTATTAATAATCGTCAAAATCTTTTACTCAAAAAATA 4777  
 QY 232 CTACCAACACATTAACCTTTTGCAGTGCAAAAAAGTA 269  
 DB 4776 ATACAACTTTCCTCCCTATTAATAACGAAAAAATA 4739

RESULT 10  
 AAS46501/c  
 ID AAS46501 standard; DNA; 9289 BP.  
 AC AAS46501;  
 XX  
 XX 18-DEC-2001 (first entry)  
 DT  
 DE Tumour suppressor gene derived chemically modified sequence #223.  
 XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;  
 KM tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
 KM cytosine methylation; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200168912-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 15-MAR-2001; 2001WO-EP002955.  
 XX  
 PR 15-MAR-2000; 2000DE-01013847.  
 PR 06-APR-2000; 2000DE-01019058.  
 PR 07-APR-2000; 2000DE-01019173.  
 PR 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2001-602752/68.  
 XX  
 XX Fragments of chemically modified genes associated with tumor suppressor  
 PT genes and oncogenes, useful in designing primers and probes for analyzing  
 PT diseases associated with cytosine methylation state e.g. cancer.  
 PS Claim 1; SEQ ID NO 223; 27P; English.

The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Ss). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Note: The sequence data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 9289 BP; 2515 A; 206 C; 2097 G; 4471 T; 0 U; 0 Other;  
 Query Match 7.0%; Score 39.6; DB 4; Length 9289;  
 Best Local Similarity 53.2%; Pred. No. 0.19; Indels 0; Gaps 0;  
 Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTGGACGCAAAAACACGCTTTTGACGCGGGCCATACATAGTCAAACT 171  
 DB 7896 AATTATTTATATTAACCAAAATTAACCTATTAATCACTACCAAAATCATTCATAAAA 7837  
 QY 172 CTACGTTTCGTAGACTATTTTACATTAATAGTCTACACCGTTGTATAGCTCAAAATCA 231  
 DB 7836 TAATTAATTTTAAACATTAATTTATTAATAATCGTCAAAATCTTTTACTCAAAAAATA 7777  
 QY 232 CTACCAACACATTAACCTTTTGCAGTGCAAAAAAGTA 269  
 DB 7776 ATACAACTTTCCTCCCTATTAATAACGAAAAAATA 7739

RESULT 11  
 ADE84121/c  
 ID ADE84121 standard; DNA; 9289 BP.  
 AC ADE84121;  
 XX  
 XX 29-JAN-2004 (first entry)  
 DT  
 DE Human lymphoid cell proliferative disorder gene derived DNA #57.  
 XX  
 KM de; lymphoid cell proliferative disorder; methylation;  
 KM methylated CpG dinucleotide; single nucleotide polymorphism; SNP;  
 KM diffuse large B-cell lymphoma; mantle cell lymphoma;  
 KM chronic lymphocytic leukemia; small lymphocytic lymphoma;  
 KM follicular lymphoma; diagnosis; prognosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO2003044226-A2.  
 XX  
 PD 30-MAY-2003.  
 XX  
 PF 25-NOV-2002; 2002WO-EP013265.  
 XX  
 PR 23-NOV-2001; 2001DE-01057491.  
 PR 28-DEC-2001; 2001DE-01064501.  
 XX  
 PA (EPIC-) EPIGENOMICS AG.  
 XX  
 PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;  
 DR WPI; 2003-457621/43.  
 XX  
 XX Detecting and differentiating between lymphoid cell proliferative  
 PT disorders comprises contacting a target nucleic acid with at least one  
 PT reagent that distinguishes between methylated and non-methylated CpG  
 PT dinucleotides.  
 XX  
 PS Claim 26; SEQ ID NO 117; 448bp; English.

The invention relates to a method of detecting and differentiating between lymphoid cell proliferative disorders associated with at least one gene and/or their regulatory regions in a subject by contacting a target nucleic acid in a biological sample obtained from the subject with at least one reagent or series of reagents that distinguish between methylated and non-methylated CpG dinucleotides within the target nucleic acid. The genes and/or their regulatory regions are preferably selected from MDRL, CSNR2B, EGR4, AR, CDK4, R82, CDC25A, Gp1b beta, MYOD1, CDH3, MYCLM, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, FOS, GSTP1, HIC-1, MGMT, MHL1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C, GSK3beta, ESR1, ADAF1, BAK1, BAX or HOUA5. Oligomers, peptide nucleic

CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences  
 CC of the genes are useful for detecting the methylation state of all the  
 CC CpG dinucleotides within one or more the sequences, or their complements,  
 CC for determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs), and for differentiating at least two of the medical  
 CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,  
 CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular  
 CC lymphoma. They are also useful for detecting of a predisposition to,  
 CC differentiation between subclasses, diagnosis, prognosis, treating and/or  
 CC monitoring of lymphoid cell proliferative disorder. This sequence  
 CC represents a nucleic acid of a pretreated genomic DNA derived from the  
 CC above mentioned genes.

SQ Sequence 9289 BP; 2515 A; 206 C; 2097 G; 4471 T; 0 U; 0 Other;

Query Match 7.0%; Score 39.6; DB 10; Length 9289;  
 Best Local Similarity 53.2%; Pred. No. 0.19;  
 Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 112 AACTTTTGGACATGCAAAAAACACGCTTTGGACGGGCCCCATCATGTGCAAACT 171  
 DB 7896 AATTATTTTATATACCAAAATAAACATTTAATCCACTACCAAAATATCATTCAAAAA 7837

QY 172 CTACGTTTGGTACATTTTACATTAATAGTCTACACCGTTGTATACGCTCAAAATCA 231  
 DB 7836 TAATTAATTTTAAACATTTATTTTAAAAATCGTCACAATTTCTTTTACTCAAAAAATA 7777

QY 232 CTACACACATTTGAACCTTTTGGACGTGCAAAAAAGTA 269  
 DB 7776 ATACAACTTTCCTCCCTCATTAATAACGAAAAATAAATA 7739

RESULT 12  
 ABZ10205/c  
 ID ABZ10205 standard; DNA; 6289 BP.

XX ABZ10205;

DT 16-JAN-2003 (first entry)

DE Haematopoietic cell proliferation disorder related DNA sequence #345.

XX Human; haematopoietic cell proliferation disorder; cytostatic;

KM Gene therapy; lymphocytic leukemia; acute myelogenous leukaemia;

XX cytosine methylation state; gene; ds.

OS Homo sapiens.

PN WO200277272-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002WO-EP003401.

PR 26-MAR-2001; 2001US-0278333P.

XX (EPiG-) EPIGENOMICS AG.

PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;  
 PI Olek A, Pienbrock C, Adorjan P, Grabs G, Liesche R, Leu B,  
 PI Lewin A, Lipschert B, Maier S, Model F, Mueller V, Otto T, Pelet C;  
 PI Schwope I, Ziebarth H;

DR WPI; 2003-018942/01.

XX Detecting and differentiating between hematopoietic cell proliferative  
 PT disorders, comprises contacting a target nucleic acid with a reagent that  
 PT distinguishes between methylated and non-methylated CpG dinucleotides.

XX Claim 28; SEQ ID NO 345; 117bp; English.

XX The present invention describes a method for detecting and  
 CC differentiating between haematopoietic cell proliferative disorders

CC associated with at least 1 gene and/or their regulatory regions in a  
 CC subject. The method comprises contacting a target nucleic acid in a  
 CC biological sample obtained from the subject with at least 1 reagent,  
 CC which distinguishes between methylated and non-methylated CpG  
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118  
 CC represent specifically claimed nucleotide sequences from the present  
 CC invention. Oligonucleotides from the present invention can be used: for  
 CC differentiation between healthy haematopoietic cells and proliferative  
 CC disorder haematopoietic cells; for differentiating between acute  
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
 CC determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder  
 CC related sequences and their complements; and as primers for the  
 CC amplification of haematopoietic cell proliferation disorder related DNA  
 CC sequences. The nucleotide sequences from the present invention can also  
 CC be used for detecting a predisposition to, differentiation between  
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of  
 CC haematopoietic cell proliferative disorders. The present method enables a  
 CC highly specific classification of haematopoietic cell proliferative  
 CC disorders allowing for improved and informed treatment of patients

SQ Sequence 6289 BP; 1680 A; 0 C; 1488 G; 3121 T; 0 U; 0 Other;

Query Match 6.7%; Score 38; DB 8; Length 6289;  
 Best Local Similarity 52.5%; Pred. No. 0.53;  
 Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 112 AACTTTTGGACATGCAAAAAACACGCTTTGGACGGGCCCCATCATGTGCAAACT 171  
 DB 4896 AATTATTTTATATACCAAAATAAACATTTAATCCACTACCAAAATATCATTCAAAAA 4837

QY 172 CTACGTTTGGTACATTTTACATTAATAGTCTACACCGTTGTATACGCTCAAAATCA 231  
 DB 4836 TAATTAATTTTAAACATTTATTTTAAAAATCATCATTAATTTCTTTTACTCAAAAAATA 4777

QY 232 CTACACACATTTGAACCTTTTGGACGTGCAAAAAAGTA 269  
 DB 4776 ATACAACTTTCCTCCCTCATTAATAACGAAAAATAAATA 4739

RESULT 13  
 ADEB84197/c  
 ID ADEB84197 standard; DNA; 9289 BP.

XX ADEB84197;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #133.

XX ds; lymphoid cell proliferative disorder; methylation;

KM methylated CpG dinucleotide; single nucleotide polymorphism; SNP;

KW diffuse large B-cell lymphoma; mantle cell lymphoma;

KW chronic lymphocytic leukemia; small lymphocytic lymphoma;

XX follicular lymphoma; diagnosis; prognosis.

OS Homo sapiens.

PN WO2003044226-A2.

PD 30-MAY-2003.

PF 25-NOV-2002; 2002WO-EP013265.

PR 23-NOV-2001; 2001DE-01057491.

PR 28-DEC-2001; 2001DE-01064501.

XX (EPiG-) EPIGENOMICS AG.

PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;

DR WPI; 2003-457621/43.

PT Detecting and differentiating between lymphoid cell proliferative  
 PT disorders comprises contacting a target nucleic acid with at least one  
 PT reagent that distinguishes between methylated and non-methylated CpG  
 PT dinucleotides.  
 PS Claim 26; SEQ ID NO 193; 448bp; English.  
 XX  
 CC The invention relates to a method of detecting and differentiating  
 CC between lymphoid cell proliferative disorders associated with at least  
 CC one gene and/or their regulatory regions in a subject by contacting a  
 CC target nucleic acid in a biological sample obtained from the subject with  
 CC at least one reagent or series of reagents that distinguish between  
 CC methylated and non-methylated CpG dinucleotides within the target nucleic  
 CC acid. The genes and/or their regulatory regions are preferably selected  
 CC from MDR1, CSNK2B, EGFR, AR, CDK4, RB2, CDC25A, GP1b beta, MYO1, CD13,  
 CC MYCL1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, CDKN2B, FOS,  
 CC GSTR1, HIC-1, MGMT, MTH1, MOS, MYC, PTEN, RBL2, TGFB2, TP73, CDKN1C,  
 CC GSK3beta, ESRI, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic  
 CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences  
 CC of the genes are useful for detecting the methylation state of all the  
 CC CpG dinucleotides within one or more the sequences, or their complements,  
 CC for determining the cytosine methylation state and/or single nucleotide  
 CC polymorphisms (SNPs), and for differentiating at least two of the medical  
 CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,  
 CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular  
 CC lymphoma. They are also useful for detecting of a predisposition to,  
 CC differentiation between subclasses, diagnosis, prognosis, treating and/or  
 CC monitoring of lymphoid cell proliferative disorder. This sequence  
 CC represents a nucleic acid of a pretreated genomic DNA derived from the  
 CC above mentioned genes.  
 XX  
 SQ Sequence 9289 BP; 2515 A; 0 C; 2097 G; 4677 T; 0 U; 0 Other;  
 Query Match 6.7%; Score 38; DB 10; Length 9289;  
 Best Local Similarity 52.5%; Pred. No. 0.63; Indels 0; Gaps 0;  
 Matches 83; Conservative 0; Mismatches 75;  
 QY 112 AACTTTTTCAGTCAAGAAAAACAGCGCTTTGACGCGGCCCATACATGATCAACT 171  
 DB 7896 AATTATTTATTAACCAATTAACATATTAATCACTACCAATATATCATTCAAAAA 7837  
 QY 172 CTACGTTTCGTAGACTATTTTACATAATAGTCTACACGCTGTATACGCTCAAAATACA 231  
 DB 7836 TAATATTTTAAACATTTATTTATTAATAATCATCAATTTCTTTTACTCAAAAAATA 7777  
 QY 232 CTACCAACATTTGAACTTTTTCAGTCAAGAAAAAGTA 269  
 DB 7776 ATCAACTTCTCTCCCTATTAATAACAAAAAATA 7739  
 RESULT 14  
 ACL35887  
 ID ACL35887 standard; cDNA; 2000 BP.  
 XX  
 AC ACL35887;  
 DT 02-JUN-2005 (first entry)  
 XX  
 DE Rice stress-regulated promoter SEQ ID NO:14450.  
 XX  
 KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;  
 XX agriculture.  
 OS Oryza sativa.  
 XX  
 PN WO2003008540-A2.  
 XX  
 PD 30-JAN-2003.  
 XX  
 PF 21-JUN-2002; 2002WO-US019668.  
 XX  
 PR 22-JUN-2001; 2001US-0300112P.  
 PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.  
 PR 21-NOV-2001; 2001US-0332132P.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,  
 PI Moughamer T, Provart N, Riecke D, Zhu T;  
 XX  
 XX WPI; 2003-248011/24.  
 DR  
 XX  
 PT New stress-responsive nucleic acid, useful for altering the  
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold  
 PT stress, salt stress or osmotic stress.  
 XX  
 PS Claim 48; SEQ ID NO 14450; 89bp; English.  
 XX  
 CC The invention relates to novel abiotic stress responsive polynucleotides  
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
 CC cells, and plants containing such polynucleotides. Also disclosed are  
 CC methods for using the polynucleotides and polypeptides to alter the  
 CC responsiveness of a plant to abiotic stress. The invention is useful in  
 CC agriculture. The nucleic acid is useful for determining whether a test  
 CC plant has been exposed to an abiotic stress condition. It is also useful  
 CC for selecting an agent that alters abiotic stress regulated  
 CC polynucleotide expression in a plant cell, and to identify a homolog or  
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
 CC molecule and the polypeptide encoded by it are useful in altering the  
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
 CC stress, osmotic stress or any of their combinations. The present sequence  
 CC is used in the exemplification of the invention  
 XX  
 SQ Sequence 2000 BP; 410 A; 329 C; 253 G; 416 T; 0 U; 592 Other;  
 Query Match 6.6%; Score 37.4; DB 11; Length 2000;  
 Best Local Similarity 16.0%; Pred. No. 0.49;  
 Matches 69; Conservative 155; Mismatches 207; Indels 0; Gaps 0;  
 QY 5 ATGATAAACAAGTATGCTGCTAATGCTGCTTCAACAACTTGTGAACTGCTTTT 64  
 DB 104 WKRYKMMWAKWATKMYKMYVATTTTTRTCAACKMKMMWGMVGTGRATTKMYAT 163  
 QY 65 CANGTTTGCAACAACACACTTATATCTCGGTGGCTCCCAACACCACTTTTGGAC 124  
 DB 164 CACTWTKWKKWRRMSAMTGMSTRTCAAMBAWMTTWSWMBRCAAMGKTGTCRNA 223  
 QY 125 TGCAAAAAACAAGCTTTTGGACGCGGCCCATATAGTACAACTTACGTTTGTAG 184  
 DB 224 GRWCATMTTWTGRTTCAGWMMKMSYMGKCAVATMTWRKRMYTAAAMTMACTYRT 283  
 QY 185 ACTATTTTACATATAATAGTCTACACCGTGTATACGCTCAATATACATACACATTG 244  
 DB 284 GMMWCMKRAAMWMAWYATTAAMWMTKAYWMCAGMBACMAWMMWMMWMMWMMWMM 343  
 QY 245 AACCTTTTGCAGTGAATAAAGTACGTGCGGACGTACGTAGCGCGGCTTGGGT 304  
 DB 344 SAAVARMAMMTTATCTAATAMMYWMSCRYKMKWMAAAAYMTMRRTMACTTAATKRMKC 403  
 QY 305 CGGTCCTGTGACGTAGCAATCATATTATCGACCGGACGAGAGTGTGCTTATCGTACA 364  
 DB 404 YGACMAAACAAYGMMTCRSTYARTWAAAGCBAAGCAAMCTTMYTKRYMTTBAAGWAA 463  
 QY 365 GAGCGCAGCTTCTGTGTGTTGTTAACCGGACGCGGACGCACTCTCTTATCGAAGAGAC 424  
 DB 464 WKCYWMMMATWTKWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMT 523  
 QY 425 GCGCCTCCATA 435  
 DB 524 WGCYSMTCAVA 534  
 RESULT 15  
 ADA71938/c  
 ID ADA71938 standard; DNA; 2000 BP.

```
XX ADA71938;
AC 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5263.
DE
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; db.
XX
XX Oryza sativa.
OS
XX MO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001MO-IB001105.
XX
XX 22-JUN-2001; 2001MO-IB001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX PI Katsigiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
XX
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
SQ
Query Match 6.4%; Score 36.2; DB 8; Length 2000;
Best Local Similarity 11.3%; Pred. No. 1.2;
Matches 42; Conservative 169; Mismatches 157; Indels 3; Gaps 1;
QY 124 CTGCAAAAAAAGCAGCTTTGGACGCGGCCCATACATAGTACAACTGTTGCTA 183
DB :||: : : : : : : : : : : : : : : : : : : : : : : : : :
430 CYGCKMWTYCSYGMKWTYMGSYKRYCYKMMYMYKGMVMMYYSAYSMWTYYY 371
QY 184 GACTATTTTACATAAATAGTCTACACCGGTGTATACCTCAATACTACACACAT 243
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
370 AKTKYKRYKRGIMSWGSKYKKYC---TWCMYKMRCYRWRKMRKKTYSKRCYC 314
QY 244 GAACTTTTTCAGTCAAAAAAGTACGTGTGGAGTCACTAGGCGGCTTATCGGG 303
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
313 RVATCYWCCCYRRSGWRSRSMRTAGMKWRSWRCSYSWKYKMKWKSKYSYMSYG 254
QY 304 TCGGCTCTGTCACTGATCAATCATCTATCGACCGGACGAGTGTCTTATCGTAC 363
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
253 WARSSTGWSAAARITKYSTSRRAKMMRACRMYSA CRYSRTSYCGSGSSSKKY 194
QY 364 AGAGCGCAGCTTCTGTGTGCTAAGCGAGCGGACGCAACTCTTATCGAAGAGA 423
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
193 MSKSGSMRMTCSNCSGCTCYGAMCMSCMSMMVWGS CGCYTRGWMKRSKYSMCKKY 134
QY 424 CGGCGCTTCATATCAGCGCGCTTATCTATGCGGTGACCGGACACAGAGCGCCGTC 483
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 133 CSCTKYCSYTGYYRYCKMYKYSYKCYCYCWYMSYMRMYMKCMCSRSSMMSGAYC 74
QY 484 CGGCTTATCGC 494
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
73 STSSTRMWSM 63
```

Search completed: June 15, 2006, 16:31:49  
Job time : 608.079 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocollateral Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 16:32:06 ; Search time 4719.79 Seconds

(without alignments)  
6682.184 Million cell updates/sec

Title: US-09-896-888a-1

Sequence: 1 catgatcataaacaatcatat.....tgctacagcagacacacatg 564

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_est7:\*  
7: gb\_est8:\*  
8: gb\_est9:\*  
9: gb\_est10:\*  
10: gb\_est11:\*  
11: gb\_est12:\*  
12: gb\_est13:\*  
13: gb\_est14:\*  
14: gb\_est15:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.2	14.4	679	1	AV398660 AV398660
2	41.4	7.3	863	13	DUA36534
3	39.6	7.0	954	12	CL077394 CH216-145
4	39	6.9	762	14	BX147578 Dancio rer
5	39	6.9	1324	12	CG751412 P045-4-E0
6	38.6	6.8	758	11	BH517933 BOCBL87TF
7	38.2	6.8	535	10	N63392 Y235H10.61
8	38	6.7	578	12	BZ902838 CH240-23B
9	38	6.7	582	12	BZ899636 CH240-16P
10	37.6	6.7	593	14	DX064229 KBrB071R0
11	37.2	6.6	425	4	CB410241 NISC nc10
12	37.2	6.6	669	8	CO384338 AGENCOURT
13	37	6.5	822	9	DN873866 nad29c08.
14	36.8	6.5	576	8	CV918268 PH005H10
15	36.8	6.5	1101	14	CNS0039X
16	36.6	6.5	432	1	AL063938 Drosophila
17	36.6	6.5	580	2	BI843287 IC59105.x
18	36.6	6.5	1032	6	CNS0GAP
19	36.4	6.5	480	4	CB484581 cclwrbho

20	36.4	6.5	857	12	CG935894	CG935894 MBEAV05TR
21	35.8	6.3	461	11	BH757407	BH757407 SALX 0561
22	35.8	6.3	655	14	AG358704	AG358704 Mus muscu
23	35.8	6.3	675	2	BG695019	BG695019 NISC_1v11
24	35.8	6.3	715	7	BE374425	BE374425 601227596
25	35.8	6.3	1049	13	CL510716	CL510716 SAHL 836
26	35.6	6.3	516	2	BM278655	BM278655 Ag tgz 65
27	35.6	6.3	613	13	CL753134	CL753134 OR_BBA012
28	35.6	6.3	626	14	AG401197	AG401197 Mus muscu
29	35.6	6.3	665	13	CL742215	CL742215 OR_BBA007
30	35.6	6.3	761	13	CM655793	CM655793 Og_BBA000
31	35.6	6.3	764	9	DN808697	DN808697 76814387
32	35.6	6.3	787	11	A0857977	A0857977 nbe0011D
33	35.6	6.3	803	9	CX336435	CX336435 JGI_XZT19
34	35.6	6.3	1080	14	CNS02APV	AL188860 Tetrarodon
35	35.4	6.3	755	11	AQ751255	AQ751255 HS_5574_B
36	35.4	6.3	835	5	CK706618	CK706618 ZP101-P00
37	35.4	6.3	997	12	CG020891	CG020891 ZMMBB055
38	35.2	6.2	579	2	CV968528	CV968528 PC063R8_1
39	35.2	6.2	815	12	CG861440	CG861440 NDL_122K2
40	35.2	6.2	1019	2	BG247625	BG247625 602359311
41	35.2	6.2	1101	14	CNS00LO0	AL068607 Drosophila
42	35	6.2	589	2	BM525801	BM525801 sak72a09.
43	35	6.2	780	14	AG394612	AG394612 Mus muscu
44	35	6.2	972	9	DN909117	DN909117 57860.3 D
45	35	6.2	988	14	CNS0067B	AL062985 Drosophila

#### ALIGNMENTS

RESULT 1  
AV398660  
LOCUS  
DEFINITION  
AV398660 Bombyx mori ovary BmNPV infected; 6 hr after inoculation  
Bombyx mori cDNA clone NV060140 T3, mRNA sequence.  
AV398660  
VERSION  
AV398660.1 GI:6902312  
KEYWORDS  
EST.  
SOURCE  
Bombyx mori (domestic silkworm)  
ORGANISM  
Bombyx mori  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Lepidoptera; Glossata; Dictyoptera;  
Bombycoidea; Bombycidae; Bombyx.

REFERENCE  
1 (bases 1 to 679)  
Mita,K., Morimoto,M., Shimada,T., Okano,K. and Maeda,S.  
AUTHORS  
Bombyx mori cDNA  
JOURNAL  
Unpublished (2000)  
COMMENT  
Contact: Mita K  
Genome Research Group  
National Institute of Agrobiological Sciences  
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan  
Tel: 81-29-838-6120  
Fax: 81-29-838-6121  
Email: kmita@ias.affrc.go.jp  
Project: 'Silkworm Genome Program in MAF, and Research for the  
Future Program in JSPS'. see 'Silkbase',  
http://www.ab.a.u-tokyo.ac.jp/silkbase/, for whole ESTdb.

#### FEATURES

source  
1..679  
/organism="Bombyx mori"  
/mol\_type="mRNA"  
/db\_xref="taxon:7091"  
/clone="NV060140"  
/cissue\_type="ovary"  
/cell\_type="BmNPV infected; 6 hr after inoculation"  
/dev\_stage="BmNPV infected; 6 hr after inoculation"  
/clone\_11b="Bombyx mori ovary BmNPV infected; 6 hr after inoculation"

#### ORIGIN

Query Match 14.4%; Score 81.2; DB 1; Length 679;  
Best Local Similarity 71.3%; Pred. No. 3.1e-14;

Matches 107; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 129 AAAAAACGCTTTTGACGCGGCGCCATACATAGTACAACTTACTTGGTACTA 188  
 DB 1 AAAAAAGTCCGCTATGTTTACATATATATACAGTACGAATCTTACAAATCGTAGCTA 60

QY 189 TTTTACATAATAGTCTACACCGCTTGATAGCTTCCAAATACACTACACACATTGAAAC 248  
 DB 61 TTTTATATAGAAATAGTCTACACCTGACATAGCTCTCAATATACTACTACATCACTA 120

QY 249 TTTTTCAGTGCACAAAAGTACGTGTCGCC 278  
 DB 121 TTTTTCATTCACAAAAGTTCATTTTGGC 150

RESULT 2  
 DU436534/c 883 bp DNA linear GSS 06-OCT-2005  
 LOCUS 1098415929593 CHORI-243 Ovis aries genomic clone CH243-284G6,  
 DEFINITION genomic survey sequence.

ACCESSION DU436534 GI:77190412  
 VERSION DU436534  
 KEYWORDS GSS.  
 SOURCE Ovis aries (sheep)  
 ORGANISM Ovis aries

REFERENCE  
 AUTHORS Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 Pecora; Bovidae; Caprinae; Ovis.  
 1 (bases 1 to 883)  
 Kikines, E., Shetty, J., de Jong, P., McEwan, J. C., Oddy, H. and  
 Cockett, N.  
 Ovine BAC End Sequences from Library CHORI-243  
 Unpublished (2004)  
 Other\_GSSes: 1098421042324  
 Contact: Ewen Kikines  
 The Institute for Genomic Research (TIGR; www.tigr.org)  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-795-7536  
 Email: ekikines@tigr.org  
 Sequences generated at the J. Craig Venter Institute Joint  
 Technology Center (JCVI/JTC; http://www.venterlinstitute.org/).  
 Original Trace: 1098415929593 Trace TI: gnl|tl|918980430  
 Insert Length: 184000 Std Error: 0.00 row: G column: 6  
 Seq Primer: T7  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..883  
 /organism="Ovis aries"  
 /mol\_type="genomic DNA"  
 /strain="Texel breed"  
 /db\_xref="taxon:9940"  
 /clone="CH243-284G6"  
 /sex="Male"  
 /cell\_type="Blood"  
 /clone\_lib="CHORI-243"  
 /notes="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 The CHORI-243 sheep (M) (Ovis aries) BAC library produced  
 by Pieter de Jong's lab at CHORI  
 http://bacpac.chori.org/library.php?id=162"

ORIGIN  
 Query Match 7.3%; Score 41.4; DB 13; Length 883;  
 Best Local Similarity 55.1%; Pred. No. 0.18; Mismatches 66; Indels 0; Gaps 0;  
 Matches 81; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAGATGATGCTAATGTTGCTTCAACAACAATTCGTGAACCTG 60  
 DB 162 CATGCTGATCAATGAGGCAATAAGATTAATTTGGTTATAGACACATTTGTGAATGTA 103

QY 61 TTTTCATGTTTGCACAAACGACCTTATATCTGCTGCTCCGCCACACCACTTTT 120  
 DB 102 TTATACCAAGCTTTTAAACACCACTTTATGTTTGGACCTTAATTCATCAACCAATGATT 43

QY 121 GCACTGCAAAAAACGCTTTTGAC 147  
 DB 42 TCAACATTAATAAAAAAGGAGTTTCTC 16

RESULT 3  
 CL077394 954 bp DNA linear GSS 31-DEC-2003  
 LOCUS CH216-145B11\_Sp6.1 CH216 Xenopus tropicalis genomic clone  
 DEFINITION CH216-145B11, genomic survey sequence.

ACCESSION CL077394 GI:40533307  
 VERSION CL077394  
 KEYWORDS GSS.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 954)  
 Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
 Mardis, E. and Wilson, R.  
 A physical map of the xenopus tropicalis genome  
 Unpublished (2003)  
 Contact: Richard K Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.wustl.edu  
 Insert Length: 175000 Std Error: 0.00  
 Seq primer: Sp6 ATTAGTGACACTATAG  
 Class: BAC ends  
 High quality sequence start: 163  
 High quality sequence stop: 226.

FEATURES  
 source Location/Qualifiers  
 1..954  
 /organism="Xenopus tropicalis"  
 /mol\_type="genomic DNA"  
 /strain="Nigerian frog"  
 /db\_xref="taxon:8364"  
 /clone="CH216-145B11"  
 /sex="male"  
 /cell\_line="Stock 248 F7A2, inbred N7"  
 /clone\_lib="CH216"  
 /notes="Vector: PTARBAC2.1; CHORI-216 Xenopus tropicalis  
 BAC library"

ORIGIN  
 Query Match 7.0%; Score 39.6; DB 12; Length 954;  
 Best Local Similarity 52.4%; Pred. No. 0.71; Mismatches 79; Indels 0; Gaps 0;  
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 103 CCACACACCAACTTTTGGACACTGCACAAAAACAGCTTTTGACGCGGCGCCATACATA 162  
 DB 330 CGCCCCCCCCCTTTTGTTTTACAAAABACACCCCTCTTCAACCCACCTTAACATA 389

QY 163 GTACAACTCTAGCTTTCGTAGACTATTTTACATAATAGTCTACACCGTTGTATCGCT 222  
 DB 390 ATGTGAAGCTCTATTTATCACTATATATATAAGGATATGACCCCTGACGTAAATC 449

QY 223 CCAATATCACTACACACATTTGAACTTTTGGACGTGCAAAAAAGT 268  
 DB 450 ATTAACATATTTACCACTGACTGCTCTCTGTAACCATATTAATT 495

RESULT 4  
 BX147578/c 762 bp DNA linear GSS 28-JAN-2003  
 LOCUS BX147578  
 DEFINITION Danio rerio genomic clone DKEX-109J14, genomic survey sequence.  
 ACCESSION BX147578  
 VERSION BX147578.1 GI:27978953  
 KEYWORDS GSS.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

/Note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
ORIGIN
Query Match          6.9%; Score 39; DB 12; Length 1324;
Best Local Similarity 48.1%; Pred. No. 1.2;
Matches 111; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 38 AACACAATTCGTGTGAAGCTGTGTTTCATGTTTGCACACAGACCTTTATCTCGGTG 97
DB 1092 AATATCAACACCTTTTAACTCTTTTCTTTTCTTTTCTTGAACCCCTTTTAAACCTCC 1151
QY 98 GCGTCCCAACCACTTTTTCGACTGCAAAAAACGCTTTTGACGCGGCGCCAT 157
DB 1152 TCCTTAGTAAACCCCAAAAGTTGAATTCATATACCGAGGCCCTTTAAATTAATCACTCA 1211
QY 158 ACATAGTACAAACCTCTACGTTTGTAGACTATTTTACATAATATGTCTACACGCTGTAT 217
DB 1212 AGAATATCCCACTAGCTTAAAGGAAATATTTCAAAAAAAATATTTCTTTTATTT 1271
QY 218 ACCTCCCAATATACACTACCAACATTAACCTTTTGACAGTGCACAAAAAGT 268
DB 1272 TGGAAAAAACAATTTTTCACCCCCCATATTTTCTTTTGCCCCAAAAAAGT 1322

RESULT 6
BHS17933          758 bp   DNA       linear   GSS 13-DEC-2001
LOCUS              BGSBL877F BGSBL Brassica oleracea genomic clone BGSBL87, genomic
DEFINITION         survey sequence.
ACCESSION           BHS17933
VERSION             BHS17933.1 GI:17726023
KEYWORDS            GSS.
SOURCE              Brassica oleracea
ORGANISM            Brassica oleracea
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE           1 (bases 1 to 758)
AUTHORS             Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
                    Uterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
                    Whole genome shotgun sequencing of Brassica oleracea and its
                    application to gene discovery and annotation in Arabidopsis
                    Genome Res. 15 (4), 487-495 (2005)
JOURNAL             PUBMED
PUBMED              15805490
COMMENT             Contact: Chris Town
                    TIGR
                    7172 Medical Center Drive, Rockville, MD 20850, USA.
                    Tel: 301-838-3523
                    Fax: 301-838-0208
                    Email: cdtown@tigr.org
                    DNA is from a doubled haploid provided by Tom Osborn.
                    Seq primer: TF
                    Class: sheared ends.
FEATURES
    source
        location/Qualifiers
            1..758
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BGSBL87"
                /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
                    genomic DNA inserted into pHOSt using BstXI linkers"
ORIGIN
Query Match          6.8%; Score 38.6; DB 11; Length 758;
Best Local Similarity 52.1%; Pred. No. 1.4;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 110 CCAACTTTTTCGACCTGCACAAAAACGCTTTTGACCGGGGCCCATATAGTACAA 169

```

Db 527 CGAGTGTGGACCTTACAGATTAATGATGACGACTCTCATCAACAATAATTATATAA 468

Qy 170 CTCTAGCTTCGAGACTATTTTACATTAATAGCTTCACCGTTGTATAGGCTCCAAATA 229

Db 467 CTCTACCAATTAATTAATAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 408

Qy 230 CACTACACACATTAAGCCTTTTTCGAGTCGCAAAAAGAGCGTGT 274

Db 407 AATTCATATCAATTAATTAATGTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTA 363

## RESULT 7

N63392 535 bp mRNA linear EST 01-MAR-1996  
LOCUS N63392  
DEFINITION Y25H10.81 Morton Fetal Cochlea Homo sapiens cDNA clone  
IMAGE:285091.3', mRNA sequence.

ACCESSION N63392  
VERSION N63392.1 GI:1211221  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 535)  
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M. F., Chiappelli, B., Chisone, S., Dietrich, N., Dubaque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucada, T., Lacy, M., Le, M., Le, N., Marids, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellendberg, K., Soares, M. B., Tan, P., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
COMMENT 8889549  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

FEATURES  
source  
1. 535  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3892208"  
/db\_xref="taxon:9606"  
/clone="IMAGE:285091"  
/tissue\_type="cochlea"  
/dev\_stage="16-22 week fetus"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/clone\_lib="Morton Fetal Cochlea"  
/note="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Reference: Genomes 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dt. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR vector. Library constructed by N. Robertson, C. Morton. -5' adaptor sequence: 5' GAATTCGGACACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

## ORIGIN

Query Match 6.8%; Score 38.2; DB 10; Length 535;  
Best Local Similarity 49.2%; Pred. No. 1.7;  
Matches 97; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 156 ATACATAGTCAACACTCTAGCTTCGAGACTATTTTACATTAATAGTCTACACCGTGT 215

Db 218 ATGTAATACCTCTCAAGCTTAGCTGTAAATAAATAAATCAACAAAGCAAAACCAATTTT 277

Qy 216 ATAGCTCCAAATACACTACACACATGAACTTTTGGAGTCAAAAAGTACGCTGT 275

Db 278 ATTTCGCCAATTTTACACAGGTTTCTTAATAATTTCTACTTATGGTAAATACAGGC 337

Qy 276 GCGAGTCAGTACGAGCCGCGCTTATCGGTCGCGCTGTACAGCTACGAATCACTATACG 335

Db 338 TCCAATCCGGTAACCAACCTTGCGATGGTTTACTTACACACTGGACACAAAGTTTACG 397

Qy 336 GACCGACGACAGTGTGT 352

Db 398 GACNCGATTAATTAATTT 414

## RESULT 8

B2902838 578 bp DNA linear GSS 12-JUN-2003  
LOCUS B2902838  
DEFINITION CH240\_23B3.TV CHORI-240 Bos taurus genomic clone CH240\_23B3, genomic survey sequence.

ACCESSION B2902838  
VERSION B2902838.1 GI:31627927  
KEYWORDS GSS.  
SOURCE Bos taurus (cattle)  
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 578)  
AUTHORS Larkin, D. M., Evers-van der Wind, A., Rebeiz, M., Schweitzer, P. A., Bachman, S., Green, C., Wright, C. L., Campos, E. J., Benson, L. D., Edwards, J., Liu, L., Osoegawa, K., Womack, J. E., de Jong, P. J. and Lewin, H. A.

TITLE A cattle-human comparative map built with cattle BAC-ends and human genome sequence  
JOURNAL Genome Res. 13 (8), 1966-1972 (2003)  
COMMENT 12902387  
Other GSSs: CH240\_23B3.TV  
Contact: Harris Lewin  
Department of Animal Sciences  
University of Illinois at Urbana Champaign  
1201 W. Gregory Dr., Urbana, IL 61801, USA  
Tel: 217 333 5998  
Fax: 217 244 5617

FEATURES  
source  
1. 578  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/strain="Breed: Hereford"  
/db\_xref="taxon:9913"  
/clone="CH240\_23B3"  
/sex="Male"  
/cell\_type="Blood"  
/clone\_lib="CHORI-240"  
/note="Vector: pTAKBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull U1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"  
Location/Qualifiers  
Class: BAC ends.

## ORIGIN

Query Match 6.7%; Score 38; DB 12; Length 578;  
Best Local Similarity 55.2%; Pred. No. 2.1;

Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 1 CAGTATGATTAACATGATGATGCTATGCTTCAACACAACTTCTGTGAACGTG 60  
 Db 252 CAGCTATCAATGAGCATTAAGATTAATGTTATAGACACATTTGTGGAATGTA 193  
 Qy 61 TTTTCATGTTGGCCAAAGACCTTTATACTCGGTGCGCTCCACACCACTTTT 120  
 Db 192 TTTTACCAAGCTTTTAAACACCACTTTATGTTGGGCTTAATCATCAGCAATGAAT 133  
 Qy 121 GCACTGCAAAAAA 134  
 Db 132 TCAGCATTAATAA 119

RESULT 9  
 BZ899636 582 bp DNA linear GSS 12-JUN-2003  
 LOCUS CH240\_16P5.TV CHORI-240 Bos taurus genomic clone CH240\_16P5,  
 DEFINITION genomic survey sequence.  
 ACCESSION BZ899636  
 VERSION BZ899636.1 GI:31624687  
 KEYWORDS GSS.  
 SOURCE Bos taurus (cattle)  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 Pecora; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 582)  
 Larkin,D.M., Everte-van der Wind,A., Rebelz,M., Schweitzer,P.A.,  
 Bachman,S., Green,C., Wright,C.L., Campos,E.J., Benson,L.D.,  
 Edwards,J., Liu,L., Osoegawa,K., Womack,J.E., de Jong,P.J. and  
 Lewin,H.A.  
 A cattle-human comparative map built with cattle BAC-ends and human  
 genome sequence  
 Genome Res. 13 (8), 1966-1972 (2003)  
 12902387  
 TITLE Contact: Harris Lewin  
 JOURNAL Department of Animal Sciences  
 PUBMED University of Illinois at Urbana Champaign  
 COMMENT 1201 W. Gregory Dr., Urbana, IL 61801, USA  
 Tel: 217 333 5998  
 Fax: 217 244 5617  
 Email: h-lewin@uiuc.edu  
 Clones are derived from the bovine BAC library CHORI-240  
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/ordering/information.htm). This work  
 was undertaken as part of the International Bovine BAC Mapping  
 Consortium (IBMC) by the University of Illinois at Urbana  
 Champaign, USA with funds provided by grant No. AG202-34480-11828  
 from USDA-CSREBS and AG99-35205-8534 from USDA/NRI (Livestock  
 Genome Sequencing Initiative)  
 Plate: 16 row: P column: 5  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source  
 1..582  
 Location/Qualifiers  
 /organism="Bos taurus"  
 /mol\_type="genomic DNA"  
 /strain="breed: Hereford"  
 /db\_xref="taxon:9913"  
 /clone="CH240\_16P5"  
 /sex="Male"  
 /cell\_type="Blood"  
 /clone\_lib="CHORI-240"  
 /note="Vector: pTAPBAC1.3; Site 1: MboI; Site 2: MboI;  
 Hereford bull LI Domino 99375; CHORI-240 Bovine BAC  
 library (Male) produced by Pieter de Jong"

ORIGIN  
 Query Match 6.7%; Score 38; DB 12; Length 582;

Best Local Similarity 55.2%; Pred. No. 2.1;

Matches 74; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 1 CAGTATGATTAACATGATGATGCTATGCTTCAACACAACTTCTGTGAACGTG 60  
 Db 265 CAGCTATCAATGAGCATTAAGATTAATGTTATAGACACATTTGTGGAATGTA 324  
 Qy 61 TTTTCATGTTGGCCAAAGACCTTTATACTCGGTGCGCTCCACACCACTTTT 120  
 Db 325 TTTTACCAAGCTTTTAAACACCACTTTATGTTGGGCTTAATCATCAGCAATGAAT 384  
 Qy 121 GCACTGCAAAAAA 134  
 Db 385 TCAGCATTAATAA 398

RESULT 10  
 DX064229 593 bp DNA linear GSS 10-JAN-2006  
 LOCUS KBRB071K09F KBrB, Brassica rapa BamHI BAC library Brassica rapa  
 DEFINITION subsp. pekinensis genomic clone KBrB071K09, genomic survey  
 sequence.  
 ACCESSION DX064229  
 VERSION DX064229.1 GI:84758525  
 KEYWORDS GSS.  
 SOURCE Brassica rapa subsp. pekinensis  
 ORGANISM Brassica rapa subsp. pekinensis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eustosida II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 593)  
 Yang,T.Y., Kwon,S.J., Kim,J.A., Kim,J.S., Lim,K.B., Jin,M.,  
 Park,J.Y., Lim,M.H., Kim,H.I., Choi,B.S., Seol,Y.J., Park,D.S.,  
 Hahn,J.H. and Park,B.S.  
 End sequence of Brassica rapa BamHI (KBrB) BAC clone  
 Unpublished (2005)  
 Contact: Beom-Seok Park  
 Brassica Genomics Team  
 National Institute of Agricultural Biotechnology  
 225 Seodun-Dong, Suwon, 441-707, Korea  
 Tel: +82-31-299-1670  
 Fax: +82-31-299-1672  
 Email: pbeom@nri.go.kr  
 BAC end sequence of Brassica rapa ssp. pekinensis BamHI BAC clone  
 KBrB071K09  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source  
 1..593  
 Location/Qualifiers  
 /organism="Brassica rapa subsp. pekinensis"  
 /mol\_type="genomic DNA"  
 /cultivar="Chifu"  
 /sub\_species="pekinensis"  
 /db\_xref="taxon:51351"  
 /clone="KBrB071K09"  
 /lab\_host="E.coli DH10B"  
 /clone\_lib="KBrB, Brassica rapa BamHI BAC library"  
 /note="Vector: pCUGIBAC1; Site 1: BamHI; Brassica rapa ssp  
 pekinensis var. Chifu BAC library (KBrB BAC) is provided  
 by Yong-Pyo Lim (CNU)."

ORIGIN  
 Query Match 6.7%; Score 37.6; DB 14; Length 593;  
 Best Local Similarity 51.2%; Pred. No. 2.8;  
 Matches 88; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 101 TCCGACACCAACTTTTGGACGCAAAAAACGCTTTGACCGGGCCCATACA 160  
 Db 39 TCCACCAATCAATTAAGTTCGCTCAAAATAGAAATCGAATATGATCAGGCCAAATA 98  
 Qy 161 TAGTACAACCTACGTTGCTAGACTATTTTACATTAATAGCTACCGCTGTATAG 220  
 Db 99 CGATTCAACTTAATATTAATTAAGAAAAACATCAAAATTTTAAAGCTTTAAAT 158

QY 221 CTCGAAATACATACACACATTTGACCTTTTTCAGCTGCAAAAATAGCT 272  
 Db 159 CTACATTGACTATACAGAAATTAATCAATTATTAACAGATTAACGACGT 210

# RESULT 11

CB410241 425 bp mRNA linear EST 24-MAR-2003  
 LOCUS NISC.nci10d09.x1 COGENE 6E MAX Homo sapiens cDNA clone IMAGE:5776553  
 DEFINITION 3', mRNA sequence.

ACCESSION CB410241  
 VERSION CB410241.1 GI:2916981  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 425)  
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 CDNA Library Preparation: CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CCGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 info@image.llnl.gov  
 Plate: LHAM12849 row: G column: 18  
 Seq primer: -21M13 forward primer (ABI).  
 Location/Qualifiers

## FEATURES

source  
 1..425  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5776553"  
 /census\_type="maxilla, pooled"  
 /dev\_stage="6 weeks postconception"  
 /lab\_host="DH10B"  
 /clone\_lib="COGENE 6E MAX"  
 /note="vector: PAMPI; CDNA primed using oligo-dT primer,  
 directionally cloned into UDG sites of PAMPI. Size  
 selected for insert sizes ranging from 0.2-1.8 kb.  
 Normalized to Cot5. Primary library, non-amplified.  
 Library constructed by M. Lovett. For more information on  
 this library, please contact R. Tidwell (Washington  
 University) or visit the COGENE website at  
 http://hg.wustl.edu/COGENE/."

## ORIGIN

Query Match 6.6%; Score 37.2; DB 4; Length 425;  
 Best Local Similarity 54.3%; Pred. No. 3.4;  
 Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 115 TTTTTCACATGCAAAAAACAGCTTTTTCACGCGGCCCATATAGTACAACTCTA 174  
 Db 66 TTTTTCGAAAAAGAAAAATTTTTCCTCGGGGTTTATTCACATGTCATATAGTGA 125  
 QY 175 CGTTGGTAGACATATTTTATATAATAGTCTACACCTTTATACGCTCCAAATACACTA 234  
 Db 126 TTTTATATGCAATATATTTGCCCCAAACTCAGCTTTTATTTTCCATTTTAAACACTA 185  
 QY 235 CCACATATGAACCTTTT 252  
 Db 186 CAATATTTACAACTGTT 203

## RESULT 12

CO384338 669 bp mRNA linear EST 30-JUN-2004  
 LOCUS AGENCOURT\_26189343 Blumberg-Cho dorsal blastopore 1lp Xenopus  
 DEFINITION laevis cDNA clone IMAGE:7297357 3', mRNA sequence.

ACCESSION CO384338  
 VERSION CO384338.1 GI:49490161  
 KEYWORDS EST.  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Xenopus.  
 1 (bases 1 to 669)  
 NIH-MGC http://mgc.ncbi.nlm.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

## JOURNAL

COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Bruce Blumberg  
 CDNA Library Preparation: B. Blumberg  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHAM15322 row: J column: 11  
 High quality sequence step: 534.  
 Location/Qualifiers

## FEATURES

source  
 1..669  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:7297357"  
 /sex="both"  
 /census\_type="dorsal blastopore 1lp"  
 /lab\_host="TOP10"  
 /clone\_lib="Blumberg-Cho dorsal blastopore 1lp"  
 /note="Organ: embryo; Vector: pluescript SK-; Site: 1:  
 EcoRI; Site: 2: XhoI; Library: prepared from 50 ug of  
 total RNA by oligo-dT priming and AMV reverse  
 transcriptase. After addition of EcoRI linker and  
 EcoRI-XhoI digestion, the cDNA was size selected by  
 chromatography on Sepharose CL-4B columns and firections  
 containing cDNAs larger than 500 bp were ligated into  
 EcoRI-XhoI-digested lambda ZAPII (Unizap-XR) and packaged  
 in vitro. Average insert size is 1.4 kb. The original  
 library contained 6 x 10<sup>6</sup> recombinants, of which 3 x 10<sup>6</sup>  
 were amplified and stored at -70 C in SM buffer containing  
 7% DMSO. 3 x 10<sup>6</sup> pfu were mass excised and the resulting  
 phageids used to infect Top10. References: Science 253,  
 196-196 and Methods in Molecular Biology 97, 555-574.  
 Additional sequences from this library have been deposited  
 under the name Xenopus laevis dorsal blastopore 1lp.  
 Library constructed by Bruce Blumberg (University of  
 California, Irvine, Department of Developmental and Cell  
 Biology)."

## ORIGIN

Query Match 6.6%; Score 37.2; DB 8; Length 669;  
 Best Local Similarity 57.9%; Pred. No. 3.8;  
 Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 153 CCATATACATATGACAACTCTACGTTTCTGTAGCATATTTTACATAAATAGTACACGCT 212  
 Db 104 CCAATTAATATAGGGGAACTTCTGTATAGTGTAAAGATACATCCAAATGATACACGT 163  
 QY 213 TGTATACGCTCCAAATACACTACACACATTTGAACTTTTTCAGTCAGCAAAAAA 266  
 Db 164 TGTATGCTCAGAAATTAATTAATCTTTTTCCTTTTTCGACGTTGAAATA 217

RESULT 13  
 DN873866 822 bp mRNA linear EST 21-APR-2005  
 LOCUS nad29c08.y1 Dog eye cornea. Unnormalized (nad) Canis familiaris  
 DEFINITION cDNA clone nad29c08 5', mRNA sequence.  
 ACCESSION DN873866  
 VERSION DN873866  
 KEYWORDS DN873866.1 GI:62843807  
 SOURCE EST.  
 ORGANISM Canis familiaris (dog)  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
 Canis.  
 1 (bases 1 to 822)  
 Wistow, G.  
 NEIBank analysis of Dog cornea  
 Unpublished (2005)  
 JOURNAL Contact: Wistow G  
 COMMENT Section on Molecular Structure and Function  
 National Eye Institute  
 6/31, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: Graeme@helix.nih.gov  
 Plate: 29 row: c column: 08  
 Seq primer: Universal M13 Reverse.  
 Location/Qualifiers  
 1..822  
 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /strain="Beagle"  
 /db\_xref="taxon:9615"  
 /clone="nad29c08"  
 /issue\_type="Cornea"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Dog eye cornea. Unnormalized (nad)"  
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted  
 from dog cornea tissue. A directionally cloned cDNA  
 library in the pCMVSPORT6 vector (Invitrogen) was  
 constructed at Bioserve Biotechnology (Laurel MD)  
 essentially following the protocols of the SuperScript  
 Plasmid System, full details of which are contained in the  
 manufacturer's instruction manual  
 (http://www.lifetechn.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 [5'-pGACTAGTCTAGATCGAGCGCGCCG(T)15-3']. cDNA was  
 cloned in Not I/Sal I sites. EST analysis was performed at  
 the NIH Intramural Sequencing Center (NISC). Analyzed data  
 available through http://neibank.nih.gov."

ORIGIN  
 Query Match 6.6%; Score 37; DB 9; Length 822;  
 Best Local Similarity 50.3%; Pred. No. 4.7;  
 Matches 91; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
 315 CACGTAAGAAATCAATATTGACCGGACGAGTGTGTTCTTATCGTACAGACGCCAGC 374  
 Db 33 CCGGACGCGCGCGGACCGGAGCCCGGAGCCGAGCCGCGCGCGGCGGCGGCGGAG 92  
 375 TTCTGTTGTTGCTAAACCGGACCGGACGCAATCTCTTATGGAACAGAGAGCCCTTCAT 434  
 Db 93 TAACTGTGCCCCGAACACCGCGCTGCGCCCTCTCCGCGCGCGCTCTCCGAGATT 152  
 435 ATGAGCGCGCGTTATCTCATGCGCGTGAACGAGACGAGAGCGCGCTCCGCTTATCGC 494  
 Db 153 CTCGGCTCCGAGCTCTCAGAGCGCGCGGAGATCTCTCTGACCGGATCTCCAGATTCCCGC 212  
 495 G 495  
 213 G 213

RESULT 14  
 CV918268/c 576 bp mRNA linear EST 25-JAN-2005  
 LOCUS PH00SH10 cycts, germinating Phytophthora infestans cDNA, mRNA  
 DEFINITION sequence.  
 ACCESSION CV918268  
 VERSION CV918268  
 KEYWORDS CV918268.1 GI:58107817  
 SOURCE EST.  
 ORGANISM Phytophthora infestans (potato late blight agent)  
 Phytophthora infestans  
 Eukaryota; stramenopiles; Oomycetes; Peronosporales; Phytophthora.  
 1 (bases 1 to 576)  
 Randall, T., Dwyer, R.A., Huitema, E., Beyer, K., Cytanich, C.,  
 Kelkar, H., Fong, A.M., Gates, K., Roberts, S., Yatzkan, E., Gaffney, T.,  
 Law, M., Testa, A., Torro-Alalibo, A., Zhang, M., Zheng, L., Mueller, E.,  
 Windass, J., Binder, A., Birch, P.R.J., Gisi, U., Govers, F., Gow, N.A.,  
 Mauch, F., van West, P., Maugh, M.E., Yu, J., Bolter, T., Kamoun, S.,  
 Lam, S.T. and Judelson, H.S.  
 Large-scale gene discovery in the oomycete phytophthora infestans  
 reveals likely components of phytopathogenicity shared with true  
 fungi  
 MOL. Plant-Microbe Interact. 18 (3), 229-243 (2005)  
 15782637  
 Contact: Judelson HS  
 Department of Plant Pathology  
 University of California  
 Webber Hall, Riverside, CA 92521, USA  
 Tel: 909 787 4199  
 Fax: 909 787 4294  
 Email: howard.judelson@ucr.edu.  
 Location/Qualifiers  
 1..576  
 /organism="Phytophthora infestans"  
 /mol\_type="mRNA"  
 /strain="88069"  
 /db\_xref="taxon:4787"  
 /sex="A1"  
 /clone\_lib="cycts, germinating"  
 /note="Vector: pSPORT1"

ORIGIN  
 Query Match 6.5%; Score 36.8; DB 8; Length 576;  
 Best Local Similarity 55.1%; Pred. No. 5;  
 Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
 8 ATAAACAATGTATGTCCTAATGCTGCTTCAACAACAATCTGTGAAGTGTTCAT 67  
 Db 494 AAAAAAAAAAANANANANNTTGTCTATGAAAACTTTTGTCAAAATTACATACTAT 435  
 68 GTTTGCCAACAGCACCTTTATATCTCGGTGGCTCCCAACCAAACTTTTGGACT 125  
 Db 434 GTGTATCAGACCAACACATCTGCGCGCTGCTGCTGCACTGATGATTTTGGACT 377

RESULT 15  
 CNS0039X/c 1101 bp DNA linear GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #  
 DEFINITION BACR08012 of RPCI-98 library from Drosophila melanogaster (fruit  
 fly) genomic survey sequence.  
 ACCESSION CNS0039X  
 VERSION AL063938  
 KEYWORDS AL063938.1 GI:4941795  
 SOURCE GSS.  
 ORGANISM Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephyridioidae; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 Genoscope.  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

## COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : segete@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see [http://www.fruitfly.org/TheBDGP/Drosophila](http://www.fruitfly.org/TheBDGP/Drosophila%20melanogaster%20BAC%20library.html)  
 melanogaster BAC library was prepared by Kazuhiro Ooegawa and  
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
 NY. The library is named RPCI-98 and was constructed by partial  
 EcoRI digestion of Drosophila DNA provided by the BDGP from the  
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
 pl and EST libraries. A more detailed description of the library  
 and how to order individual BAC clones, the entire library, or  
 filters for hybridization from the BACPAC Resource Center can be  
 found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source  
 1. 1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR08012"  
 /clone\_lib="RPCI-98"  
 /note="end : T7"

## ORIGIN

Query Match 6.5%; Score 36.8; DB 14; Length 1101;  
 Best Local Similarity 35.9%; Pred. No. 5.8;  
 Matches 61; Conservative 35; Mismatches 74; Indels 0; Gaps 0;  
 QY 73 CCAACAAGACGCTTTATCTGGTGGCTCCACGACCAACTTTTGCATGCAAAA 132  
 Db 1068 CATTMTATTTACACACATATATCTCTCTTCTTATACAAATAATWACTATWCACTMTATA 1009  
 QY 133 AACACGCTTTGACGCGGCGCCATACATAGTACAACTCTACGTTTGTAGACTATTTT 192  
 Db 1008 CMCACYYAWCACAAMCMTCCCTCTCTATATAATCAATCTAATAATCAMAAYTACMAH 949  
 QY 193 ACATTAATAGTACACCGCTTGTATAGGCTCCAATACACTACACACAT 242  
 Db 948 MAMAAMAWATATWCMAYHMTTTCACACACACTATACWCMWACYACWT 899

Search completed: June 15, 2006, 18:15:40  
 Job time : 4724.79 secs



Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	36.4	6.5	565	3	US-09-621-976-15639	Sequence 15639, A	
	2	34.4	6.1	855	3	US-09-148-545-89	Sequence 89, Appl	
	3	34.4	6.1	855	3	US-09-621-011-89	Sequence 89, Appl	
C	4	33.8	6.0	3083	3	US-08-693-308-1	Sequence 1, Appl	
C	5	33.2	5.9	902	3	US-09-370-767-2513	Sequence 2513, A	
	6	33.2	5.9	5003	3	US-09-370-767-9827	Sequence 9827, Ap	
	7	32.2	5.7	601	3	US-09-549-016-103819	Sequence 103819,	
C	8	32.2	5.7	601	3	US-09-549-016-103820	Sequence 103820,	
C	9	32.2	5.7	601	3	US-09-549-016-103821	Sequence 103821,	
	10	32.2	5.7	601	3	US-09-549-016-103863	Sequence 103863,	
	11	32.2	5.7	601	3	US-09-549-016-103864	Sequence 103864,	
C	12	32.2	5.7	601	3	US-09-549-016-103865	Sequence 103865,	
	13	32.2	5.7	601	3	US-09-549-016-160230	Sequence 160230,	
	14	32.2	5.7	601	3	US-09-549-016-160231	Sequence 160231,	
C	15	32.2	5.7	601	3	US-09-549-016-160232	Sequence 160232,	
	16	32.2	5.7	601	3	US-09-549-016-160274	Sequence 160274,	
	17	32.2	5.7	601	3	US-09-549-016-160275	Sequence 160275,	
C	18	32.2	5.7	601	3	US-09-549-016-160276	Sequence 160276,	
	19	32.2	5.7	78649	3	US-09-549-016-14619	Sequence 14619, A	
	20	32.2	5.7	78649	3	US-09-549-016-14620	Sequence 14620, A	
C	21	32.2	5.7	78649	3	US-09-549-016-16227	Sequence 16227, A	
	22	32.2	5.7	78649	3	US-09-549-016-16228	Sequence 16228, A	
	23	31.4	5.6	54452	3	US-09-549-016-12642	Sequence 12642, A	

C	24	31.4	5.6	54452	3	US-09-949-016-13003	Sequence 13003, A
	25	31.2	5.5	34125	3	US-09-782-378A-25	Sequence 25, Appl1
C	26	31.2	5.5	35775	3	US-09-949-016-13505	Sequence 13505, A
	27	31	5.5	549	3	US-09-902-540-6433	Sequence 6433, Ap
C	28	31	5.5	601	3	US-09-949-016-86304	Sequence 86304, A
	29	31	5.5	601	3	US-09-949-016-86305	Sequence 86305, A
C	30	31	5.5	601	3	US-09-949-016-86306	Sequence 86306, A
	31	31	5.5	601	3	US-09-949-016-86307	Sequence 86307, A
C	32	31	5.5	601	3	US-09-949-016-86308	Sequence 86308, A
	33	31	5.5	601	3	US-09-949-016-86309	Sequence 86309, A
C	34	31	5.5	601	3	US-09-949-016-86310	Sequence 86310, A
	35	31	5.5	7218	2	US-08-232-463-14	Sequence 14, Appl1
C	36	31	5.5	12865	3	US-09-902-540-1048	Sequence 1048, Ap
	37	31	5.5	74790	3	US-09-949-016-15321	Sequence 15321, A
C	38	31	5.5	422592	3	US-09-949-016-11182	Sequence 14182, A
	39	30.8	5.5	1701	3	US-09-949-016-80113	Sequence 80113, A
C	40	30.8	5.5	1710	3	US-09-252-991A-12652	Sequence 12652, A
	41	30.8	5.5	64313	3	US-09-949-016-11597	Sequence 11597, A
C	42	30.8	5.5	70313	3	US-09-949-016-16064	Sequence 16064, A
	43	30.8	5.5	271134	3	US-09-949-016-12705	Sequence 12705, A
C	44	30.8	5.5	305931	3	US-09-949-016-17550	Sequence 17550, A
	45	30.6	5.4	4396	3	US-09-821-736-1	Sequence 1, Appl1

## ALIGNMENTS

```

RESULT 1
US-09-621-976-15639/C
: Sequence 15639, Application US/09621976
: Patent No. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSET 054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent .pm
: SEQ ID NO 15639
: LENGTH: 505
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-621-976-15639

```

Query Match	6.5%	Score 36.4	DB 3	Length 505

Matches 40; Conservative 159; Mismatches 165; Indels 0; Gaps 0;

Oy	12	GCAGAAAAACGCGTTTTCAGCGGGGCCATACATAGTACAAATCTACGTTTCGTAGA	18
Db	364	RNAAAGAGCTCGMTSYSGSKMTGRGSKMTKRKMMTYSGGMMTTSYKCTKTKKTYGMSK	30
Oy	186	CTATTTTACATAATAGTCTACACCGCTGTATACGCTCAATACATACATCCACATTTGA	24
Db	304	KTRMTCTSRKTYMMNMSGMARSMKSMASRSMASMAACGMSASAYRRARRMYGARBSM	24
Oy	246	ACCTTTTTCAGTCGAAAAAGTACGTGTCGCACTCAGTAGCGCGCTTATCGGTC	30
Db	244	RAGAGMMRARGKTRARGKSSMMRSMKMSMRMSAKAKMRMRMMSCRMMSYCMGSKCM	18
Oy	306	GGCTCGTCGACGTACGAATCATTTATTCGACCGGAGAGTGTTCTTATCGTACAG	36
Db	184	SCGCTCAKMRAYARAKRYAASSMGKYMGCRCYAKCAKMYGYKRSRSTGSGMKYRR	12
Oy	366	GAGCCAGCTCTCTGTGTCTTAACCGGACCGGACGCAACTCTTATTCGGAAGAGACG	42
Db	124	RKMYVMKTYMMNMSWMCYRMGAMYGMSARAYRMTASMSACMGCSMMNMSMMNMCWR	65
Oy	426	CGCTTCATATACGCCGCGCTTATCTATGCGCGTGAACCGACACGAGCGCCCGTCC	48

Db 64 SRYRQWMSGKWCYSCGYCCSACRMGYCMTBRKMSWYSRKSRLMCRYSMSAYRYSK 5  
Qy 486 GCTT 489  
Db 4 RTRT 1

RESULT 2  
US-09-148-545-89  
Sequence 89, Application US/09148545  
Patent No. 6590075  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 70 Human Secreted Proteins  
FILE REFERENCE: P2001P1  
CURRENT APPLICATION NUMBER: US/09/148, 545  
EARLIER FILING DATE: 1998-09-04  
EARLIER APPLICATION NUMBER: PCT/US98/04482  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,161  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631

```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 89
; LENGTH: 855

```

```

Query Match 6.1%; Score 34.4; DB 3; Length 855;
Best Local Similarity 55.6%; Pred. No. 0.41;
Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

Qy 439 GCGCGCGTTATCTCATGCGCGTGACCGGACAGAGCGCGCGCTTATCGCGCTT 498
Db 12 GCGCGACCCGAGCTCAGCTCGCGCCACCCACCAAGTTCCAGTGGCGCGACGAGTGGCTT 71
Qy 499 ATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAGCATCTGTACAGCGA 555

```

```

Db 72 ATGCGTCCCTCACTGCGGCTGCGACAGAGNACTTGACTGACGAGATGCGACGCA 128
RESULT 3
US-09-621-011-89
; Sequence 89, Application US/09621011
; Patent No. 6878687
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/621,011
; CURRENT FILING DATE: 2000-07-20
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 89
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (103)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (767)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (831)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-621-011-89

```

```

Query Match 6.1%; Score 34.4; DB 3; Length 855;
Best Local Similarity 55.6%; Pred. No. 0.41;
Matches 65; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

Qy 439 GCGCGCGTTATCTCATGCGCGTGACCGGACAGAGCGCGCGCTTATCGCGCTT 498
Db 12 GCGCGACCCGAGCTCAGCTCGCGCCACCCACCAAGTTCCAGTGGCGCGACGAGTGGCTT 71
Qy 499 ATAAATACAGCCCGCAACGATCTGTAAACACAGTTGAACAGCATCTGTACAGCGA 555
Db 72 ATGCGTCCCTCACTGCGGCTGCGACAGAGNACTTGACTGACGAGATGCGACGCA 128
RESULT 4
US-08-693-308-1/c
; Sequence 1, Application US/08693308
; Patent No. 6447996
; GENERAL INFORMATION:
; APPLICANT: HABERT-ORTOLI, Estelle
; APPLICANT: AMIRANOFF, Brigitte
; APPLICANT: LOQUET, Isabelle
; TITLE OF INVENTION: GALANIN RECEPTOR, NUCLEIC ACIDS,
; TITLE OF INVENTION: TRANSFORMED CELLS AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,308
; FILING DATE:
; CLASSIFICATION: 536

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 94/01808  
FILING DATE: 17-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/FR95/00172  
FILING DATE: 14-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: 38,619  
REFERENCE/DOCKET NUMBER: ST94008-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3083 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 787..1836  
US-08-693-308-1

Query Match 6.0%; Score 33.8; DB 3; Length 3083;  
Best Local Similarity 54.4%; Pred. No. 1.3;

Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 10 AACCATGTATGCTGCTTAATGCTTCAACACAAATTCGTGACTGTCTTCTATGT 69  
DB 2939 AAATAATGTTGGTGAATAATGTCATTAACATCATTAACAATTAAGAAAAAGCATGC 2880  
70 TTGCGCAACAGACCTTATACCTGCGGCTCCGCCACCACTTTTGGACCTGCAA 129  
DB 2879 ATTGTRCAAGATAGATACCCAGATATTTTCAGTCCCATGATTTTGTGAATGTCAT 2820  
QY 130 AAAAA 134  
DB 2819 CTADA 2815

RESULT 5  
US-09-270-767-25113/c  
Sequence 25113, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 25113  
LENGTH: 902  
TYPE: DNA  
ORGANISM: Drosophila melanogaster  
US-09-270-767-25113

Query Match 5.9%; Score 33.2; DB 3; Length 902;  
Best Local Similarity 56.4%; Pred. No. 1.1;  
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 205 TACACGTTGTATACGCTCCAAATATACACACATTAACCTTTTGCAGTGCAGAAA 264  
DB 440 TAACCTGATGTAGCATCTCCAGATACATAGAGCTCATAGAACTGTTGAAAAGGAAAT 381  
QY 265 AAGTAGTGTGGCAGTCACTAGAGCCGCTTATCGGTCGGTCTCTGT 314

DB 380 ATTCCGATCTGTGTCTCAGAGCGGCTGTGAACCCTGCTGCTT 331

RESULT 6  
US-09-270-767-9827/c  
Sequence 9827, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 9827  
LENGTH: 5003  
TYPE: DNA  
ORGANISM: Drosophila melanogaster  
US-09-270-767-9827

Query Match 5.9%; Score 33.2; DB 3; Length 5003;  
Best Local Similarity 56.4%; Pred. No. 2.7;  
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 205 TACACGTTGTATACGCTCCAAATATACACACATTAACCTTTTGCAGTGCAGAAA 264  
DB 1212 TAACCTGATGTAGCATCTCCAGATACATTAAGCTGTATGAAGGAAAT 1153  
QY 265 AAGTAGTGTGGCAGTCACTAGAGCGGCTTATCGGTCGGTCTCTGT 314  
DB 1152 ATTCCGATCTGTGTCTCAGAGCGGCTGTGAACCCTGCTGCTT 1103

RESULT 7  
US-09-949-016-103819  
Sequence 103819, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 103819  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-103819

Query Match 5.7%; Score 32.2; DB 3; Length 601;  
Best Local Similarity 56.0%; Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7 GATAAACAATGATAGTGTCTTAATGTTGCTTCAACAACAATTCGTGAACTGTGTTCA 66  
DB 463 GCTAAACAATGATATTTGAGACTTTTCTTCAGAAATTAATGATGAAATGTAATTAA 522  
QY 67 TGTGGCAACAAGACCTTTTATCTCGGTGCGCTCCACAGCAACT 115  
DB 523 TCTATCAATATGCTCTTTCTTAATCAATGTTCCCATCTCAGATT 571

RESULT 8  
US-09-949-016-103820  
; Sequence 103820, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103820  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-103820

Query Match  
Best Local Similarity 56.0%; Score 32.2; DB 3; Length 601;  
Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACAATGATGCTGCTAATGCTGCTTCAACAACAATCTGTGAAGTGTTCCTTCA 66  
Db 133 GCTAAACATGATGATTTTGAGACTTTCTTCTTCAGAAATTATGATCTGAAGTGTATTAA 192

Qy 67 TGTTCGCAACAGCACCTTTATCTGCGTGGCTCCCAACCAACT 115  
Db 193 TCTATCAGATGATGCTCTTTCTTCAATCATGATGTTCCCAATCTCAGATT 241

RESULT 9  
US-09-949-016-103821  
; Sequence 103821, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103821  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-103821

Query Match  
Best Local Similarity 56.0%; Score 32.2; DB 3; Length 601;  
Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACAATGATGCTGCTAATGCTGCTTCAACAACAATCTGTGAAGTGTTCCTTCA 66  
Db 102 GCTAAACATGATGATTTTGAGACTTTCTTCTTCAGAAATTATGATCTGAAGTGTATTAA 161

Qy 67 TGTTCGCAACAGCACCTTTATCTGCGTGGCTCCCAACCAACT 115  
Db 193 TCTATCAGATGATGCTCTTTCTTCAATCATGATGTTCCCAATCTCAGATT 241

Db 162 TCTATCAGATGATGCTCTTTCTTCAATCATGATGTTCCCAATCTCAGATT 210

RESULT 10  
US-09-949-016-103863  
; Sequence 103863, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103863  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-103863

Query Match  
Best Local Similarity 56.0%; Score 32.2; DB 3; Length 601;  
Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACAATGATGCTGCTAATGCTGCTTCAACAACAATCTGTGAAGTGTTCCTTCA 66  
Db 463 GCTAAACATGATGATTTTGAGACTTTCTTCTTCAGAAATTATGATCTGAAGTGTATTAA 522

Qy 67 TGTTCGCAACAGCACCTTTATCTGCGTGGCTCCCAACCAACT 115  
Db 523 TCTATCAGATGATGCTCTTTCTTCAATCATGATGTTCCCAATCTCAGATT 571

RESULT 11  
US-09-949-016-103864  
; Sequence 103864, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103864  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-103864

Query Match  
Best Local Similarity 56.0%; Score 32.2; DB 3; Length 601;  
Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACAATGATGCTGCTAATGCTGCTTCAACAACAATCTGTGAAGTGTTCCTTCA 66  
Db 133 GCTAAACATGATGATTTTGAGACTTTCTTCTTCAGAAATTATGATCTGAAGTGTATTAA 192

Qy 67 TGTTCGCAACAGACCTTTATCTGGTGGCTCCCAACCACT 115  
Db 193 TCTATCAGATATGCTCTTTCTAATCCATGTTCCCATCTCAGATT 241

## RESULT 12

US-09-949-016-103865  
; Sequence 103865, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103865  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-103865

Query Match 5.7%; Score 32.2; DB 3; Length 601;  
Best Local Similarity 56.0%; Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACATATGATGTGCTAATGTTCTTCACACAACTTGTGAAGTGTTC 66  
Db 102 GCTAAACATATGATATTTGAGACTTTTCTTCAGAAATATGATCTGAAGTGTATTTAA 161  
Qy 67 TGTTCGCAACAGACCTTTATCTGGTGGCTCCCAACCACT 115  
Db 162 TCTATCAGATATGCTCTTTCTAATCCATGTTCCCATCTCAGATT 210

## RESULT 13

US-09-949-016-160230  
; Sequence 160230, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 160230  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-160230

Query Match 5.7%; Score 32.2; DB 3; Length 601;  
Best Local Similarity 56.0%; Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACATATGATGTGCTAATGTTCTTCACACAACTTGTGAAGTGTTC 66  
Db 463 GCTAAACATATGATATTTGAGACTTTTCTTCAGAAATATGATCTGAAGTGTATTTAA 522  
Qy 67 TGTTCGCAACAGACCTTTATCTGGTGGCTCCCAACCACT 115  
Db 523 TCTATCAGATATGCTCTTTCTAATCCATGTTCCCATCTCAGATT 571

## RESULT 14

US-09-949-016-160231  
; Sequence 160231, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 160231  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-160231

Query Match 5.7%; Score 32.2; DB 3; Length 601;  
Best Local Similarity 56.0%; Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 7 GATTAACATATGATGTGCTAATGTTCTTCACACAACTTGTGAAGTGTTC 66  
Db 133 GCTAAACATATGATATTTGAGACTTTTCTTCAGAAATATGATCTGAAGTGTATTTAA 192  
Qy 67 TGTTCGCAACAGACCTTTATCTGGTGGCTCCCAACCACT 115  
Db 193 TCTATCAGATATGCTCTTTCTAATCCATGTTCCCATCTCAGATT 241

## RESULT 15

US-09-949-016-160232  
; Sequence 160232, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 160232  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-160232

Query Match 5.7%; Score 32.2; DB 3; Length 601;

Best Local Similarity 56.0%; Pred. No. 1.9;  
Matches 61; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY	7	GATTAACAAATGTATGCTAATGCTTGCACACAAATTCTGTGAACTGTGTTTCA	66
Db	102	GCTAAACATAGATATTTGAGACTTTTCTTCAGGAATTATGATCTGAAAGTTGATTTAA	161
QY	67	TGTTTGCCACAAACCACTTTATATCTGGTGGCCTCCCCACCACT	115
Db	162	TCTATCAGAAATAGCTCTTTCTAATCAATTGTTCCCAATCTCAGATT	210

Search completed: June 15, 2006, 17:18:09  
Job time : 195.328 secs

**THIS PAGE BLANK (u&rw)**



Result No.	Query				ID	Description
	Score	Match	Length	DB		
1	564	100.0	564	3	US-09-896-8888-1	Sequence 1, Appli
2	540	97.2	550	9	US-10-622-088-126	Sequence 126, App
3	548	97.2	273	7	US-10-595-074-60	Sequence 60, Appl
4	548	97.2	273	9	US-10-846-911-60	Sequence 60, Appl
5	548	97.2	273	10	US-10-622-088-89	Sequence 60, Appl
6	546.4	96.9	5038	9	US-10-939-107-60	Sequence 89, Appl
7	462	81.9	462	9	US-09-896-8888-14	Sequence 14, Appl
8	462	81.9	147	9	US-10-622-088-127	Sequence 127, App
9	62	11.0	325	9	US-10-622-088-14	Sequence 149, App
10	62	11.0	325	9	US-09-896-8888-15	Sequence 15, Appl
11	57	10.1	92	3	US-09-896-8888-15	Sequence 15, Appl
12	42.6	7.6	88	3	US-10-473-126-199	Sequence 199, App
13	39.6	7.0	6289	9	US-10-473-126-199	Sequence 222, App
14	39.6	6.7	9289	8	US-10-473-126-199	Sequence 344, App
15	35.4	6.3	6289	9	US-10-473-126-345	Sequence 1528, App
16	35.4	6.3	966	8	US-10-398-221-1548	Sequence 3152, App
17	34.4	6.1	855	3	US-10-398-221-3152	Sequence 819, Appl
					US-09-881-876-89	

18	-34.4	6.1	855	3	US-09-148-545-89	Sequence 89, Appl
19	34.4	6.1	855	10	US-10-979-111-89	Sequence 89, Appl
20	34.2	6.1	436	9	US-10-425-115-104158	Sequence 104158, Appl
21	34	6.0	18977	13	US-11-097-143-26311	Sequence 26311, A
22	33.8	6.0	600	13	US-11-060-756-2178	Sequence 2178, Appl
23	33.8	6.0	600	13	US-11-060-756-2179	Sequence 2179, Appl
24	33.8	6.0	600	13	US-11-060-756-6450	Sequence 6450, Appl
25	33.8	6.0	600	13	US-11-060-756-6451	Sequence 6451, Appl
26	33.8	6.0	930	4	US-09-925-065A-77884	Sequence 27884, A
27	33.8	6.0	930	5	US-09-925-065A-77884	Sequence 27884, A
28	33.8	6.0	930	12	US-10-301-480-129121	Sequence 129121, A
29	33.8	6.0	930	12	US-10-301-480-142530	Sequence 742530, A
30	33.8	6.0	126	4	US-09-925-065A-68947	Sequence 68947, A
31	33.8	6.0	126	5	US-09-925-065A-68947	Sequence 68947, A
32	33.8	6.0	126	12	US-10-301-480-170186	Sequence 170186, A
33	33.8	6.0	126	12	US-10-301-480-783595	Sequence 783595, A
34	33.8	6.0	2117	6	US-10-087-192-1463	Sequence 1463, Appl
35	33.8	6.0	3056	6	US-10-225-367A-125	Sequence 125, Appl
36	33.8	6.0	3083	7	US-10-166-568-1	Sequence 1, Appl
37	33.8	6.0	39344	6	US-10-087-192-1462	Sequence 1462, Appl
38	33.2	5.9	573	4	US-09-925-065A-559402	Sequence 559402, Appl
39	33.2	5.9	573	5	US-09-925-065A-559402	Sequence 559402, Appl
40	33.2	5.9	621	8	US-10-437-963-77138	Sequence 77138, A
41	33	5.9	14429	7	US-10-311-455-2215	Sequence 2215, Appl
42	33	5.9	14429	8	US-10-433-793-127	Sequence 127, Appl
43	32.8	5.8	426	10	US-10-779-543-9174	Sequence 9174, Appl
44	32.8	5.8	454	3	US-09-770-444-580	Sequence 580, Appl
45	32.8	5.8	2000	3	US-09-938-842A-5148	Sequence 5148, Appl

## ALIGNMENTS

```

RESULT 1
US-09-896-888A-1
: Sequence 1, Application US/09896888A
: Patent No. US2002016723A1
: GENERAL INFORMATION:
: APPLICANT: The University of British Columbia
: TITLE OF INVENTION: Insect Expression Vectors
: FILE REFERENCE: 80021-44
: CURRENT APPLICATION NUMBER: US/09/896,888A
: PRIORITY FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US/09/048,911
: PRIORITY FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/049,946
: PRIORITY FILING DATE: 1997-03-27
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: Patencin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 564
: TYPE: DNA
: ORGANISM: Oryza pseudotibugata
: US-09-896-888A-1

```

Query Match	100.0%	Score 564	DB 3	length 564
Best Local Similarity	100.0%	Pred. No. 4	8e-15	
Matches 564	0	Mismatches	0	Gaps 0

Qy	1	CAGGATGATGAAACATGATGAGTGAATGTTGCTTCAACAAACATCTGTTGAACGTG	60
		2	
Db	1	CATGATGATTAACAATGTAATGAGTCTAAATGTTGCTTCAACAAACATCTGTTGAACGTG	60
Qy	61	TTTTCAATGTTGCCAACAAGCCTTTTAATCTCGGTGCGCTCCGACCCACCAACTTTT	120
Db	61	TTTTCAATGTTGCCAACAAGCCTTTTAATCTCGGTGCGCTCCGACCCACCAACTTTT	120
Qy	121	GCACATGATGAAACACGCTTTTGACGCGGGCCCATCATAGTACAAACCTTCAGTTTC	180
Db	121	GCACATGATGAAACACGCTTTTGACGCGGGCCCATCATAGTACAAACCTTCAGTTTC	180
Qy	181	GTGACATATTTTACATAAATAGCTACACCGTTGTATGCGTCCAAATACATACACAC	240
Db	181	GTGACATATTTTACATAAATAGCTACACCGTTGTATGCGTCCAAATACATACACAC	240

```
Db 181 GTAGACTATTATTAATAATAGTCTACACCGTGTATACGCTCCAAATATACACAC 240
Qy 241 ATTGAACCTTTTTCAGTGCAGAAAAAGTACGTGTGCGGCACTCAGTATAGCGGCTTATC 300
Db 241 ATTGAACCTTTTTCAGTGCAGAAAAAGTACGTGTGCGGCACTCAGTATAGCGGCTTATC 300
Qy 301 GGGTCGCGTCTGTGCACTACGATGATCACTTATGAGACCGGAGAGTGTGCTTATCGT 360
Db 301 GGGTCGCGTCTGTGCACTACGATGATCACTTATGAGACCGGAGAGTGTGCTTATCGT 360
Qy 361 GACAGAGCGCAGGCTTCTGTGTGCTTAAACCGACCGGAGCGAACTCTTATGGAACA 420
Db 361 GACAGAGCGCAGGCTTCTGTGTGCTTAAACCGACCGGAGCGAACTCTTATGGAACA 420
Qy 421 GGAAGCGGCTTCCATATCAGCGCGGCTTATCTATGCGCGTACCGGACAGAGCGGCCC 480
Db 421 GGAAGCGGCTTCCATATCAGCGCGGCTTATCTATGCGCGTACCGGACAGAGCGGCCC 480
Qy 481 GTCCCGCTTATCGCGGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Db 481 GTCCCGCTTATCGCGGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Qy 541 CATCTGTTACAGCAGCAACATG 564
Db 541 CATCTGTTACAGCAGCAACATG 564
```

```
RESULT 2
US-10-622-088-126
; Sequence 126, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Opi2 promoter sequence
US-10-622-088-126
```

```
Query Match 97.2%; Score 548; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 9.2e-170;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CATGATGATAAACAATGATGCTATGCTTCAACAACAATCTGTGAACGTG 60
Db 5 CATGATGATAAACAATGATGCTATGCTTCAACAACAATCTGTGAACGTG 64
Qy 61 TTTTCATGTTGGCAACAGACCTTTATCTCGGTGCGCTCCCAACCAACTTTT 120
```

```
Db 65 TTTTCATGTTGGCAACAGACCTTTATCTCGGTGCGCTCCCAACCAACTTTT 124
Qy 121 GCATCGCAAAAAAACAAGCTTTTGGACGCGGCGCCATATAGTATCAAACTGACGTTTC 180
Db 125 GCATCGCAAAAAAACAAGCTTTTGGACGCGGCGCCATATAGTATCAAACTGACGTTTC 184
Qy 181 GTAGACTATTATTAATAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 240
Db 185 GTAGACTATTATTAATAATAGTCTACACCGTGTATACGCTCCAAATACATACACAC 244
Qy 241 ATTGAACCTTTTTCAGTGCAGAAAAAGTACGTGTGCGGCACTCAGTATAGCGGCTTATC 300
Db 245 ATTGAACCTTTTTCAGTGCAGAAAAAGTACGTGTGCGGCACTCAGTATAGCGGCTTATC 304
Qy 301 GGGTCGCGTCTGTGCACTACGATGATCACTTATGAGACCGGAGAGTGTGCTTATCGT 360
Db 305 GGGTCGCGTCTGTGCACTACGATGATCACTTATGAGACCGGAGAGTGTGCTTATCGT 364
Qy 361 GACAGAGCGCAGGCTTCTGTGTGCTTAAACCGACCGGAGCGAACTCTTATGGAACA 420
Db 365 GACAGAGCGCAGGCTTCTGTGTGCTTAAACCGACCGGAGCGAACTCTTATGGAACA 424
Qy 421 GGAAGCGGCTTCCATATCAGCGCGGCTTATCTATGCGCGTACCGGACAGAGCGGCCC 480
Db 425 GGAAGCGGCTTCCATATCAGCGCGGCTTATCTATGCGCGTACCGGACAGAGCGGCCC 484
Qy 481 GTCCCGCTTATCGCGGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Db 485 GTCCCGCTTATCGCGGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 544
Qy 541 CATCTGTT 548
Db 545 CATCTGTT 552
```

```
RESULT 3
US-10-295-074-60
; Sequence 60, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p220p2f expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
```

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0,

```

Db          545 CATCTGTT 552

RESULT 4
US-10-846-911-60
Sequence 60, Application US/10846911
Publication No. US20040258660A1
GENERAL INFORMATION:
APPLICANT: KLYSNER, Steen
APPLICANT: NIELSEN, Finn Stausholm
APPLICANT: BRATT, Tomas
APPLICANT: VOLDBORG, Bjorn
APPLICANT: MOURITSEN, Soren
TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
FILE REFERENCE: 674542-2018
CURRENT FILING DATE: 2004-05-14
PRIOR APPLICATION NUMBER: PCT/DK02/00764
PRIOR FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: 60/331,575
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: PA 2001 01702
PRIOR FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: Patentin version 3.1
SEQ ID NO 60
LENGTH: 2773
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: p22Op2F expression vector for insect cells
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (561)..(566)
OTHER INFORMATION: HindiII site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (573)..(578)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (586)..(591)
OTHER INFORMATION: EcoRI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (593)..(598)
OTHER INFORMATION: BamHI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (625)..(630)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (629)..(634)
OTHER INFORMATION: ClaI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (1156)..(1161)
OTHER INFORMATION: ApaLI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2128)..(2133)
OTHER INFORMATION: PciI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2204)..(2209)
OTHER INFORMATION: NcoI site
FEATURE:
NAME/KEY: misc_recomb
LOCATION: (2284)..(2289)
OTHER INFORMATION: AvalI site
FEATURE:
NAME/KEY: misc_recomb

```

LOCATION: (2294)..(2299)  
 OTHER INFORMATION: Aval, SmaI, and XmaI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2551)..(2556)  
 OTHER INFORMATION: ApalI site  
 US-10-846-911-60

Query Match 97.2%; Score 548; DB 9; Length 2773;

Best Local Similarity 100.0%; Pred. No. 2e-169; Mismatches 0; Indels 0; Gaps 0;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATATATATGCTATGTTGCTTCAACAACAACTTCTGTGAAGTGTG 60  
 DB 5 CATGATGATTAACAATATATGCTATGTTGCTTCAACAACAACTTCTGTGAAGTGTG 64  
 QY 61 TTTTCATGTTTGGCAACAAGACCTTTATATCTGGTGCCCTCCCAACCAACTTTT 120  
 DB 65 TTTTCATGTTTGGCAACAAGACCTTTATATCTGGTGCCCTCCCAACCAACTTTT 124  
 QY 121 GCACGTGAAAAAACAAGCTTTTGCACGGGGCCCATACATAGTACAACTTACGTTTC 180  
 DB 125 GCACGTGAAAAAACAAGCTTTTGCACGGGGCCCATACATAGTACAACTTACGTTTC 184  
 QY 181 GTAGACTATTTTACATTAATAGTCTACACCGTTGTATACGCTCAATAACACTACACAC 240  
 DB 185 GTAGACTATTTTACATTAATAGTCTACACCGTTGTATACGCTCAATAACACTACACAC 244  
 QY 241 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTGCGACGTACAGTACGCGGCTTATC 300  
 DB 245 ATTGAACCTTTTTCAGTGCAGTCAAAAAAGTACGTGCGACGTACAGTACGCGGCTTATC 304  
 QY 301 GGGTGGGCTCTGTCAGTGCAGTCAATATATGGAACCGGACGAGTGTCTATACGT 360  
 DB 305 GGGTGGGCTCTGTCAGTGCAGTCAATATATGGAACCGGACGAGTGTCTATACGT 364  
 QY 361 GACAGACGCGCAAGCTTCTGTGTGTCTAACGCGACCGGACGCAACTCTTATCGAACA 420  
 DB 365 GACAGACGCGCAAGCTTCTGTGTGTCTAACGCGACCGGACGCAACTCTTATCGAACA 424  
 QY 421 GACGCGGCTTCCATATACGCGCGGTTATCTCATGCGGCTGACCGGACGAGCGCCC 480  
 DB 425 GACGCGGCTTCCATATACGCGCGGTTATCTCATGCGGCTGACCGGACGAGCGCCC 484  
 QY 481 GTCCCGGTTATTCGGCCTATTAATAACGCCGCAACATCTGTGAACAGTGAACAG 540  
 DB 485 GTCCCGGTTATTCGGCCTATTAATAACGCCGCAACATCTGTGAACAGTGAACAG 544  
 QY 541 CATCTGTT 548  
 DB 545 CATCTGTT 552

RESULT 5  
 US-10-939-107-60

Sequence 60, Application US/10939107  
 Publication No. US20050180947A1

GENERAL INFORMATION:

APPLICANT: Pedersen, Hans Rudolf

APPLICANT: Ebert, Bjarte

APPLICANT: Pedersen, Louise Henriette

APPLICANT: Rasmussen, Peter Birk

TITLE OF INVENTION: Novel Application of Vaccination Against TNF-alpha

FILE REFERENCE: 674542-2020

CURRENT APPLICATION NUMBER: US/10/939,107

PRIOR FILING DATE: 2004-09-10

PRIOR APPLICATION NUMBER: PCT/DK03/00147

PRIOR FILING DATE: 2003-03-11

PRIOR APPLICATION NUMBER: 60/363,128

NUMBER OF SEQ ID NOS: 60

SOFTWARE: PatentIn version 3.1

SEQ ID NO 60

LENGTH: 2773  
 TYPE: DNA  
 ORGANISM: Artificial sequence  
 FEATURE:  
 OTHER INFORMATION: p2ZOP2F expression vector for insect cells  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (561)..(566)  
 OTHER INFORMATION: HindIII site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (573)..(578)  
 OTHER INFORMATION: Aval site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (586)..(591)  
 OTHER INFORMATION: EcoRI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (593)..(598)  
 OTHER INFORMATION: BamHI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (625)..(630)  
 OTHER INFORMATION: ClaI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (629)..(634)  
 OTHER INFORMATION: ClaI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (1156)..(1161)  
 OTHER INFORMATION: ApalI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2128)..(2133)  
 OTHER INFORMATION: PstI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2204)..(2209)  
 OTHER INFORMATION: NcoI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2284)..(2289)  
 OTHER INFORMATION: Aval site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2294)..(2299)  
 OTHER INFORMATION: Aval, SmaI, and XmaI site  
 FEATURE:  
 NAME/KEY: misc\_recomb  
 LOCATION: (2551)..(2556)  
 OTHER INFORMATION: ApalI site  
 US-10-939-107-60

Query Match 97.2%; Score 548; DB 10; Length 2773;

Best Local Similarity 100.0%; Pred. No. 2e-169; Mismatches 0; Indels 0; Gaps 0;

Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGATGATTAACAATATATGCTATGTTGCTTCAACAACAACTTCTGTGAAGTGTG 60  
 DB 5 CATGATGATTAACAATATATGCTATGTTGCTTCAACAACAACTTCTGTGAAGTGTG 64  
 QY 61 TTTTCATGTTTGGCAACAAGACCTTTATATCTGGTGCCCTCCCAACCAACTTTT 120  
 DB 65 TTTTCATGTTTGGCAACAAGACCTTTATATCTGGTGCCCTCCCAACCAACTTTT 124  
 QY 121 GCACGTGAAAAAACAAGCTTTTGCACGGGGCCCATACATAGTACAACTTACGTTTC 180  
 DB 125 GCACGTGAAAAAACAAGCTTTTGCACGGGGCCCATACATAGTACAACTTACGTTTC 184  
 QY 181 GTAGACTATTTTACATTAATAGTCTACACCGTTGTATACGCTCAATAACACTACACAC 240

```

Db      185 GTAGACTATTTTACATAAATAGCTACACCGTGTATACGCTCCAAATACACTACACAC 244
Qy      241 ATTGAACCTTTTGTGAGTGAAGAAAAAGTACGTCGAGTACAGTACGCGGCTTATC 300
Db      245 ATTGAACCTTTTGTGAGTGAAGAAAAAGTACGTCGAGTACAGTACGCGGCTTATC 304
Qy      301 GGGTCGGCTCTGTACAGTACGAATACATTATCGSACCGGACGAGTGTCTTATCGT 360
Db      305 GGGTCGGCTCTGTACAGTACGAATACATTATCGSACCGGACGAGTGTCTTATCGT 364
Qy      361 GACAGAGCGCAGCTTCTGTGTGTCTAACCGGACCGGACGCAACTCTTATCGAACA 420
Db      365 GACAGAGCGCAGCTTCTGTGTGTCTAACCGGACCGGACGCAACTCTTATCGAACA 424
Qy      421 GGAAGCGCTCTCATATACAGCGCGGCTTATCTATGCGCGTACCGGACGAGCGGCC 480
Db      425 GGAAGCGCTCTCATATACAGCGCGGCTTATCTATGCGCGTACCGGACGAGCGGCC 484
Qy      481 GTCCCGCTTATCGCGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Db      485 GTCCCGCTTATCGCGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 544
Qy      541 CATCTGTT 548
Db      545 CATCTGTT 552

```

## RESULT 6

```

US-10-622-088-89
; Sequence 89, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Frank, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PIB/V5-His-DEST
US-10-622-088-89

```

```

Query Match      96.9%; Score 546.4; DB 9; Length 5038;
Best Local Similarity 99.8%; Pred. No. 9.2e-169;
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 CAGGATATTAACAAGTATGCTAATGTTGCTTAACAACAATCTGTGAACGTG 60
Db      1 CAGGATATTAACAAGTATGCTAATGTTGCTTAACAACAATCTGTGAACGTG 60
Qy      61 TTTTCATGTTTGCACACAGCAGCTTATTAATCTGGTGGCTCCACCAACTTTT 120

```

```

Db      61 TTTTCATGTTTGCACACAGCAGCTTATTAATCTGGTGGCTCCACCAACTTTT 120
Qy      121 GCACGTCAAAAAACACGCTTTTGGACGCGGCCCATATCATATGATCAAACTTACGTTTC 180
Db      121 GCACGTCAAAAAACACGCTTTTGGACGCGGCCCATATCATATGATCAAACTTACGTTTC 180
Qy      181 GTAGACTATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACACAC 240
Db      181 GTAGACTATTTTACATAAATAGTCTACACCGTGTATACGCTCCAAATACACTACACAC 240
Qy      241 ATTGAACCTTTTGTGAGTGAAGAAAAAGTACGTCGAGTACAGTACGCGGCTTATC 300
Db      241 ATTGAACCTTTTGTGAGTGAAGAAAAAGTACGTCGAGTACAGTACGCGGCTTATC 300
Qy      301 GGGTCGGCTCTGTACAGTACGAATACATTATCGSACCGGACGAGTGTCTTATCGT 360
Db      301 GGGTCGGCTCTGTACAGTACGAATACATTATCGSACCGGACGAGTGTCTTATCGT 360
Qy      361 GACAGAGCGCAGCTTCTGTGTGTCTAACCGGACCGGACGCAACTCTTATCGAACA 420
Db      361 GACAGAGCGCAGCTTCTGTGTGTCTAACCGGACCGGACGCAACTCTTATCGAACA 420
Qy      421 GGAAGCGCTCTCATATACAGCGCGGCTTATCTATGCGCGTACCGGACGAGCGGCC 480
Db      421 GGAAGCGCTCTCATATACAGCGCGGCTTATCTATGCGCGTACCGGACGAGCGGCC 480
Qy      481 GTCCCGCTTATCGCGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Db      481 GTCCCGCTTATCGCGCTTATTAATATACAGCCGCAACGATCGTAAACAGTTGAACAG 540
Qy      541 CATCTGTT 548
Db      545 CATCTGTT 548

```

## RESULT 7

```

US-09-896-888a-14
; Sequence 14, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Promoter
; OTHER INFORMATION: sequence of the Opmpv 1e2 gene
US-09-896-888a-14

```

```

Query Match      81.9%; Score 462; DB 3; Length 462;
Best Local Similarity 100.0%; Pred. No. 2.1e-141;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      102 CCCACACCAACTTTTGTGACTGCAAAAAACAGCTTTTGGACGCGGCCCATATCAT 161
Db      1 CCCACACCAACTTTTGTGACTGCAAAAAACAGCTTTTGGACGCGGCCCATATCAT 160
Qy      162 AGTACAACTCTACGTTTGTGAGATATTTTACATTAATAGTCTACACCGTTGTATACGC 221
Db      61 AGTACAACTCTACGTTTGTGAGATATTTTACATTAATAGTCTACACCGTTGTATACGC 220

```



PRIOR APPLICATION NUMBER: US/09/048,911  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/049,946  
PRIOR FILING DATE: 1997-03-27  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 16  
LENGTH: 92  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fragment of  
OTHER INFORMATION: promoter sequence of the AcMNPV ien gene  
US-09-896-888a-16

Query Match 10.1%; Score 57; DB 3; Length 92;  
Best Local Similarity 77.5%; Pred. No. 5.9e-08;  
Matches 69; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 186 CTAATTTACATAATAGTCTACACCGTGTATACGCTCCCAATACACTACACACATGA 245  
Db 1 CTAATTTATGAAATGACTCTACACTGTACATACGCTCCCAATATACTACTACACTATCA 60

Qy 246 ACCTTTTGCAGTGCAGAAAAGTACGTGT 274  
Db 61 ACTTTTTCATTAACAAAAGTTCAATTT 89

RESULT 11  
US-09-896-888a-15  
Sequence 15, Application US/09896888A  
Patent No. US20020116723A1  
GENERAL INFORMATION:  
APPLICANT: The University of British Columbia  
TITLE OF INVENTION: Insect Expression Vectors  
FILE REFERENCE: 80021-44  
CURRENT APPLICATION NUMBER: US/09/896,888A  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US/09/048,911  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/049,946  
PRIOR FILING DATE: 1997-03-27  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 15  
LENGTH: 88  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Fragment of  
OTHER INFORMATION: promoter sequence of the AcMNPV ien gene  
US-09-896-888a-15

Query Match 7.6%; Score 42.6; DB 3; Length 88;  
Best Local Similarity 74.0%; Pred. No. 0.0033;  
Matches 54; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 112 AACTTTTGCACGTGCAAAAAACGCTTTTGCACGGGGCCCATACATAGTACAACT 171  
Db 16 AATTTTTCGATGCAAAAAAGTCACTTTGCTGACACACTCATATACAGTACATCT 75

Qy 172 CTACGTTGGTAG 184  
Db 76 CTACAAATCGTAG 88

RESULT 12  
US-10-473-126-199/c  
Sequence 199, Application US/10473126  
Publication No. US20040234973A1  
GENERAL INFORMATION:  
APPLICANT: Epigenomics AG  
TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell

TITLE OF INVENTION: proliferative disorders  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/473,126  
CURRENT FILING DATE: 2003-09-26  
NUMBER OF SEQ ID NOS: 1258  
SEQ ID NO 199  
LENGTH: 6289  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-473-126-199

Query Match 7.0%; Score 39.6; DB 9; Length 6289;  
Best Local Similarity 53.2%; Pred. No. 0.27;  
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 112 AACTTTTGCACGTGCAAAAAACGCTTTTGCACGGGGCCCATACATAGTACAACT 171  
Db 4896 AATTTATTTATATACCAATATAACTATTATTAATCCACTACCAATATCATTCAAAA 4837

Qy 172 CTACGTTGCTAGCTATTTTACATATAATAGTCTACACCGCTGTATACGCTCAATACA 231  
Db 4836 TAATAATTTTAAACATATTTATTAATAAATCGTCACAATCTTTTACTCAAAAAATA 4777

Qy 232 CTACCAACATGCAACCTTTTGCACGTGCAAAAAATA 269  
Db 4776 ATACAACTTCTCTCTCTATATAAAGCAAAAAATA 4739

RESULT 13  
US-10-221-714A-223/c  
Sequence 223, Application US/10221714A  
Publication No. US20040048254A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with  
TITLE OF INVENTION: tumor suppressor genes and oncogenes  
FILE REFERENCE: 5013.1005  
CURRENT APPLICATION NUMBER: US/10/221,714A  
CURRENT FILING DATE: 2003-01-21  
PRIOR APPLICATION NUMBER: PCT/EP01/02955  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: DE 10013847.0  
PRIOR FILING DATE: 2000-03-15  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 540  
SEQ ID NO 223  
LENGTH: 9289  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-221-714A-223

Query Match 7.0%; Score 39.6; DB 8; Length 9289;  
Best Local Similarity 53.2%; Pred. No. 0.32;  
Matches 84; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 112 AACTTTTGCACGTGCAAAAAACGCTTTTGCACGGGGCCCATACATAGTACAACT 171  
Db 7896 AATTTATTTATATACCAATATAACTATTATTAATCCACTACCAATATCATTCAAAA 7837

Qy 172 CTACGTTGCTAGCTATTTTACATATAATAGTCTACACCGCTGTATACGCTCAAAATACA 231

Db 7836 TAAATATTTTAAACATTAATTAATAAATGCTCAATTTCTTTACTCAAAAAATA 7777  
QY 232 CTACACACATGTGAACCTTTTTCAGTGCAGCAAAAAAGTA 269  
Db 7776 ATACAACTTCTCTCTCCCTATTAATAAGCAAAAAATA 7739

## RESULT 14

US-10-473-126-345/c  
; Sequence 345, Application US/10473126  
; Publication No. US20040234973A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell  
; TITLE OF INVENTION: Proliferative disorders  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/473,126  
; CURRENT FILING DATE: 2003-09-26  
; NUMBER OF SEQ ID NOS: 1258  
; SEQ ID NO 345  
; LENGTH: 6289  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-473-126-345

Query Match 6.7%; Score 38; DB 9; Length 6289;  
Best Local Similarity 52.5%; Pred. No. 0.9;  
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 112 AACCTTTTGACAGCAAAACACGCTTTTGACAGCGGCGCCATACATAGTACAACT 171  
Db 4896 AATTTATTTATATACCAATTAACCTATTAATCCACTACCAATATCATTCAAAA 4837  
QY 172 CTACGTTTCGTAGCTATTTTACATAAATAGCTACACCGTGTATACGCTCAATACA 231  
Db 4836 TAAATATTTTAAACATTTATTAATAAATCATCAATTTCTTTACTCAAAAAATA 4777  
QY 232 CTACACACATGTGAACCTTTTTCAGTGCAGCAAAAAAGTA 269  
Db 4776 ATACAACTTCTCTCTCCCTATTAATAAGCAAAAAATA 4739

## RESULT 15

US-10-398-221-1528/c  
; Sequence 1528, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederik  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: *Listeria innocua*, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; CURRENT FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1528  
; LENGTH: 986  
; TYPE: DNA  
; ORGANISM: *Listeria monocytogenes*-4B  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)..(end)  
; OTHER INFORMATION: n can be any nucleotide a,g,c or t/u  
US-10-398-221-1528

Query Match 6.3%; Score 35.4; DB 8; Length 986;

Best Local Similarity 56.4%; Pred. No. 2.6;  
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 155 CATACATAGTACAAACTTACGTTTGTGACATTTTACATTAATAGTCTACACCGTG 214  
Db 318 CAGATTGAAAAACATTTATCGTTGCCAGAAAAATTAATAATCAAGACACTCTGTG 259  
QY 215 TATACGCTCCAAATFACATACACACATTTGACATTTTTCAGTGCAGCAAAAAAGTACG 271  
Db 258 CAGAACATTCATTAATAGTTACATCCATTGCTCAATTTTTCGCTGCTGAGAGAG 202

Search completed: June 15, 2006, 18:02:42  
Job time : 1907.88 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 17:21:03 ; Search time 62.413 Seconds

(without alignments)  
2127.111 Million cell updates/sec

Title: US-09-896-888a-1

Perfect score: 564

Sequence: 1 catgatgataacacatgtat.....tgttaacgacacacatg 564

Scoring table: IDENTITY\_NUC

Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA New:  
1: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
3: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	32.8	5.8	1216	6	US-10-953-349-868 Sequence 868, App
C 2	32.8	5.8	3119	6	US-10-449-902-9093 Sequence 9093, App
C 3	31.6	5.6	1127	6	US-10-449-902-2563 Sequence 2563, App
C 4	31.6	5.6	1128	6	US-10-449-902-19810 Sequence 19810, App
C 5	31.6	5.6	7833	6	US-10-517-441-350 Sequence 350, App
C 6	31.6	5.6	7833	6	US-10-517-441-350 Sequence 624, App
C 7	31.4	5.6	7873	6	US-10-485-397-7 Sequence 7, Appl
C 8	31.4	5.6	7943	6	US-10-485-397-8 Sequence 8, Appl
C 9	30.8	5.5	1006	6	US-10-449-902-24096 Sequence 24096, App
C 10	30.6	5.4	9353	6	US-10-517-441-462 Sequence 462, App
C 11	30.6	5.4	9353	6	US-10-517-441-736 Sequence 736, App
C 12	30.2	5.4	2019	6	US-10-449-902-22046 Sequence 22046, App
C 13	30.2	5.4	2137	6	US-10-449-902-9697 Sequence 9697, App
C 14	30.0	5.3	1114	6	US-10-953-349-32119 Sequence 32119, App
C 15	30.0	5.3	1237	6	US-10-511-837-430 Sequence 430, App
C 16	30.0	5.3	2214	6	US-10-449-902-25649 Sequence 25649, App
C 17	30.0	5.3	2586	6	US-10-449-902-24690 Sequence 24690, App
C 18	30.0	5.3	2618	6	US-10-449-902-16689 Sequence 16689, App
C 19	30.0	5.3	3107	6	US-10-517-441-401 Sequence 401, App
C 20	29.6	5.2	3501	6	US-10-517-441-499 Sequence 499, App
C 21	29.6	5.2	1266	6	US-10-953-349-3657 Sequence 3657, App
C 22	29.6	5.2	3283	6	US-10-517-441-678 Sequence 678, App
C 23	29.4	5.2	3331	6	US-10-953-349-8124 Sequence 8124, App
C 24	29.2	5.2	2538	7	US-11-217-529-218 Sequence 218, App
C 25	28.4	5.0	838	6	US-10-449-902-9307 Sequence 9307, App

C 26	28.4	5.0	1456	6	US-10-449-902-12236 Sequence 12236, App
C 27	28.4	5.0	2108	6	US-10-953-349-5230 Sequence 5230, App
C 28	28.4	5.0	4611	7	US-11-267-871-717 Sequence 717, App
C 29	28.4	5.0	4644	1	US-09-484-331-19 Sequence 19, App
C 30	28.4	5.0	5247	1	US-09-484-331-20 Sequence 20, App
C 31	28.4	5.0	5264	6	US-10-485-397-10 Sequence 10, App
C 32	28.4	5.0	5314	1	US-09-484-331-27 Sequence 27, App
C 33	28.4	5.0	5408	6	US-10-485-397-9 Sequence 9, App
C 34	28.4	5.0	5715	7	US-11-258-392-4 Sequence 4, App
C 35	28.4	5.0	6233	7	US-11-296-119-10 Sequence 10, App
C 36	28.4	5.0	9737	1	US-09-484-331-22 Sequence 22, App
C 37	28.4	5.0	9737	1	US-09-484-331-23 Sequence 23, App
C 38	28.4	5.0	9737	1	US-09-484-331-28 Sequence 28, App
C 39	28.4	5.0	9871	1	US-09-484-331-24 Sequence 24, App
C 40	28.4	5.0	10060	1	US-09-484-331-25 Sequence 25, App
C 41	28.4	5.0	11772	7	US-11-246-405-17 Sequence 17, App
C 42	28.2	5.0	687	7	US-11-242-317-38 Sequence 38, App
C 43	28.2	5.0	1688	6	US-10-449-902-10918 Sequence 10918, App
C 44	28.2	5.0	2416	7	US-11-289-102-76 Sequence 76, App
C 45	28.2	5.0	3214	6	US-10-449-902-18041 Sequence 18041, App

#### ALIGNMENTS

```
RESULT 1
US-10-953-349-868/c
; Sequence 868, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 868
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-953-349-868

Query Match      5.8%; Score 32.8; DB 6; Length 1216;
Best Local Similarity 53.0%; Pred. No. 0.17;
Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY      5 ATGATAACAGTATGTCCTAATGTCCTCAACAACTCTGTGAACGTGTTT 64
DB      292 AACCCAAACGCGCTTATACCTCAATCCTCACTACACCTTCCGAAGACTGTCATC 233
QY      65 CATGTTGCCAACAAACACCTTATATCTGCGTCCGCCACCACTTTTGCAC 124
DB      232 CATTAAGCTCAATCCGCACTTTGTTGATATCTTACAGTACCAACTTGTGATC 173
QY      125 TGCACAAAACA 136
DB      172 GTACAAATACA 161

RESULT 2
US-10-449-902-9093/c
; Sequence 9093, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USBS THEROP
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
```

```

; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9093
; LENGTH: 3119
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: AK106441
; DATABASE ENTRY DATE: 2002-08-28
; US-10-449-902-9093
```

```

Query Match          5.6%; Score 32.8; DB 6; Length 3119;
Best Local Similarity 67.6%; Pred. No. 0.29;
Matches 46; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```

QY 190 TTTACATTAATGCTACACCGCTGTATGCTTCCAAATACATCACACATTTGAACCT 249
DB 1385 TTTGAAGAAATGTCACCTCATTTGATTCCTCAAAATCCATTTGAAGATTGACCA 1326
QY 250 TTTTGCGAG 257
DB 1325 TTTTCCGAG 1318
```

```

RESULT 3
US-10-449-902-2563/c
; Sequence 2563, Application US/10449902
; Publication No. US20060123505A1
```

```

; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2563
; LENGTH: 1127
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: AK104298
; DATABASE ENTRY DATE: 2002-08-28
; US-10-449-902-2563
```

```

Query Match          5.6%; Score 31.6; DB 6; Length 1127;
Best Local Similarity 52.2%; Pred. No. 0.43;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```

QY 347 TGTGTCTTATCGACAGACGCGCAGCTTCTGTGTTGCTAACCGACGCGACGAC 406
DB 218 TGTGTACTGCGCCGCGCCGCGCGCGCTGTCTCCGACATCGCGGCGCCGAC 159
QY 407 TCTTATCGAAGACGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGGTGACG 466
DB 158 CTCGACACCAATCTCTCGCGCGCGCGCTCTCTCCGACGAGATCGACGCGGCGCTCG 99
QY 467 GACACGAGGCGCC 480
DB 98 GTGACGCTGAGGCC 85
```

```

RESULT 4
US-10-449-902-19810/c
; Sequence 19810, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 19810
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: AK070153
; DATABASE ENTRY DATE: 2001-12-06
; US-10-449-902-19810
```

```

Query Match          5.6%; Score 31.6; DB 6; Length 1128;
Best Local Similarity 52.2%; Pred. No. 0.43;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```

QY 347 TGTGTCTTATCGACAGACGCGCAGCTTCTGTGTTGCTAACCGACGCGACCCAC 406
DB 219 TGTGTACTGCGCCGCGCCGCGCGCGCTGTCTCCGACATCGCGGCGCCGAC 160
QY 407 TCTTATCGAAGACGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGGTGACG 466
DB 159 CTCGACACCAATCTCTCGCGCGCGCGCTCTCTCCGACGAGATCGACGCGGCGCTCG 100
QY 467 GACACGAGGCGCC 480
DB 99 GTGACGCTGAGGCC 86
```

```

RESULT 5
US-10-517-441-350/c
; Sequence 350, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBEC, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamara
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEPLER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell]
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010861
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
```

Query Match	5.6%	Score	31.6	DB	6	Length	7833
Best Local	50.0%	Pred.	No.	1.2			
Matches	79	Conservative	0	Mismatches	79	Indels	0
				Gaps			0

```

1 ORGANISM: Artificial
2
3 FEATURE:
4
5 OTHER INFORMATION: pcgpstpoosectis is a circular plasmid DNA

```

US-10-485-397-8

Query Match 5.6%; Score 31.4; DB 6; Length 7943;  
Best Local Similarity 45.9%; Pred. No. 1.5;  
Matches 107; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

Qy 321 CGATCATATTATCGACCGGACGAGTGTCTTATCGAGACGACGCGAGCTTCG 380  
Db 4319 CCGAGCCGATGAGCGGCGCTGCTCTGCTTGCAGTTATTCAGGACCGGACAGGTGCG 4260  
Qy 381 TGTGTACCGGACCGGACGCAACTCTTATCGAACAAGACGCGCTTCATATCAGC 440  
Db 4259 TCTTGAACAAAAGAACCGGGGCGCCCTGCGCTGACAGCCGGAACAGCGCGCATCAAGC 4200  
Qy 441 CGGCGCTTATCTCTGCGGCTGACCGGACGAGCGCCGCTCCGCTTATCGCGCTAT 500  
Db 4199 AGCCGATGTCTGTGTGGCCAGTCATAGCCGAATACCTCTCCACCCAGCGCGCGAG 4140  
Qy 501 AATATACAGCCGCAACGATGTGTAACACAGTTGAACAGCATCTGTTACAGC 553  
Db 4139 AACCTGCTGATCCATCTTGTTCATTAATCTAGTTCACTAGTTCTTTACAGC 4087

RESULT 9

US-10-449-902-24096/c  
Sequence 24096, Application US/10449902  
Publication No. US20060123505A1

GENERAL INFORMATION:  
APPLICANT: National Institute of Agrobiological Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
FILE REFERENCE: MOA-A02051-US  
CURRENT APPLICATION NUMBER: US/10/449, 902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 24096  
LENGTH: 1006  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK099538  
DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-24096

Query Match 5.5%; Score 30.8; DB 6; Length 1006;  
Best Local Similarity 51.4%; Pred. No. 0.77;  
Matches 71; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 160 ATAGTCAAACTCACTGTTTGTAGACTATTTTACATTAATAGCTACACCGTGTATAC 219  
Db 898 AGATGACAACTCCATCTCTGTTGGAGAAAAAAGACATCTAGACTGGTTAACA 839  
Qy 220 GCTTCAAAATACACTACACACATTTGAGTGTGCAAAAAAGTACGTGTCCGA 279  
Db 838 GGTGCTACTACTACTATGATGAGACAAATTAACAAGAAATGTTATGTGATGCG 779  
Qy 280 GTACAGTAGCGCGGCTT 297  
Db 778 TGCTCATGTGTGCTT 761

RESULT 10

US-10-517-441-462/c  
Sequence 462, Application US/10517441  
Publication No. US20060121467A1  
GENERAL INFORMATION:

APPLICANT: FOEKENS, John  
APPLICANT: HARBECK, Nadia  
APPLICANT: KOENIG, Thomas  
APPLICANT: MAIER, Sabine  
APPLICANT: MARTENS, John  
APPLICANT: MODEL, Fabian  
APPLICANT: NIMMICH, Inko  
APPLICANT: RUJAN, Tamas  
APPLICANT: SCHMITT, Armin  
APPLICANT: SCHMITT, Manfred  
APPLICANT: LOOK, Maxime P.  
APPLICANT: MARX, Almut  
APPLICANT: HOEFLER, Heinz  
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
FILE REFERENCE: 47675-93  
CURRENT APPLICATION NUMBER: US/10/517, 441  
CURRENT FILING DATE: 2004-12-11  
PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
PRIOR FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: DE 10317955.0  
PRIOR FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: DE 1030096.8  
PRIOR FILING DATE: 2003-01-07  
PRIOR APPLICATION NUMBER: DE 10245779.4  
PRIOR FILING DATE: 2002-10-01  
NUMBER OF SEQ ID NOS: 2147  
SEQ ID NO 462  
LENGTH: 9353  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-462

Query Match 5.4%; Score 30.6; DB 6; Length 9353;  
Best Local Similarity 49.1%; Pred. No. 3;  
Matches 81; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 10 AATCATGTATGTGCTATATGTTGCTTCAACAAGATTCTGTGAACGTGTTTATGT 69  
Db 8082 ATAAATTTATCTTACTAATTTCTAAATTAACAACCTTACTTCCAAATCTCTTAAAT 8023  
Qy 70 TTGCAACAAGACCTTTATCTCGGTGCGTCCGACCAACCTTTTGGACTGCA 129  
Db 8022 TCTAAATCAACCCATCAAACTTTTAACTTAAACATCAAAAAATTAATCTTAACTTTA 7963  
Qy 130 AAAACACGCTTTTGCACGCGGCGCCATACATAGTACAACTCTA 174  
Db 7962 ATAAAACTTACTTAACCAACCACTTAATATACAAACAACTA 7918

RESULT 11

US-10-517-441-736/c  
Sequence 736, Application US/10517441  
Publication No. US20060121467A1

GENERAL INFORMATION:  
APPLICANT: FOEKENS, John  
APPLICANT: HARBECK, Nadia  
APPLICANT: KOENIG, Thomas  
APPLICANT: MAIER, Sabine  
APPLICANT: MARTENS, John  
APPLICANT: MODEL, Fabian  
APPLICANT: NIMMICH, Inko  
APPLICANT: RUJAN, Tamas  
APPLICANT: SCHMITT, Armin  
APPLICANT: SCHMITT, Manfred  
APPLICANT: LOOK, Maxime P.  
APPLICANT: MARX, Almut  
APPLICANT: HOEFLER, Heinz  
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
FILE REFERENCE: 47675-93

CURRENT APPLICATION NUMBER: US/10/517,441  
CURRENT FILING DATE: 2004-12-11  
PRIORITY APPLICATION NUMBER: PCT/EP2003/010881  
PRIOR FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: DE 10317955.0  
PRIOR FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: DE 10300096.8  
PRIOR FILING DATE: 2003-01-07  
PRIOR APPLICATION NUMBER: DE 10245779.4  
PRIOR FILING DATE: 2002-10-01  
NUMBER OF SEQ ID NOS: 2147  
SEQ ID NO 736  
LENGTH: 9353  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-736

Query Match 5.4%; Score 30.6; DB 6; Length 9353;  
Best Local Similarity 49.1%; Pred. No. 3;  
Matches 81; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 10 AACAAATGATGGTATGCTTATGCTTCAACAACATTTGTTGAAGTGTTCATGT 69  
DB 8082 ATAAATTTTCTTACTTAATTAATTAACACCTTACTTTCAAAATTTCTTAAAT 8023  
QY 70 TTGCCAACAGACCTTTATCTGCTGCGCTCCCAACCACTTTTTCAGTCAA 129  
DB 8022 TCTAAATCAAAACCATCAAACTTTTACCTCAAAACATCAAAATCTAATTTTA 7963  
QY 130 AAAAAGCGCTTTTGACGGGGCCCATACATAGTACAACTCTA 174  
DB 7962 ATAAAAAATTAATAACACACACTATATACAAACACACTA 7918

## RESULT 12

US-10-449-902-22046  
Sequence 22046, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449,902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22046  
LENGTH: 2019  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK072256  
DATABASE ENTRY DATE: 2001-12-06  
US-10-449-902-22046

Query Match 5.4%; Score 30.2; DB 6; Length 2019;  
Best Local Similarity 55.1%; Pred. No. 1.8;  
Matches 59; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 388 AACCGACGCGGACGCACTCTTATGGAACAGAGCGCCCTTCATATACGCGCGGT 447  
DB 1562 AAGGTATCATGACGAGCGGCTCCCTCCCGACGCGCGCCGCGCGGT 1621  
QY 448 TATCTATGCGGTGACCGGACAGAGGCGCGCTCCCTTATCGC 494

DB 1622 GACCTATCCGCGCCTCTCTGTAAGAGCCCAAGCGGATCCG 1668

## RESULT 13

US-10-449-902-9697  
Sequence 9697, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449,902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9697  
LENGTH: 2127  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK107045  
DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-9697

Query Match 5.4%; Score 30.2; DB 6; Length 2127;  
Best Local Similarity 55.1%; Pred. No. 1.9;  
Matches 59; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 388 AACCGACGCGGACGCACTCTTATGGAACAGAGCGCCCTTCATATACGCGCGGT 447  
DB 1543 AAGGTATCATGACGAGCGGCTCCCTCCCGACGCGCGCGCGGT 1602  
QY 448 TATCTATGCGGTGACCGGACAGAGGCGCGCTCCCTTATCGC 494  
DB 1603 GACCTATCCGCGCCTCTCTGTAAGAGCCCAAGCGGATCCG 1649

## RESULT 14

US-10-953-349-32119/c  
Sequence 32119, Application US/10953349  
Publication No. US20060107345A1  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
TITLE OF INVENTION: ENCODED THERBY  
FILE REFERENCE: 2750-1579PUS2  
CURRENT APPLICATION NUMBER: US/10/953,349  
CURRENT FILING DATE: 2004-09-30  
NUMBER OF SEQ ID NOS: 40252  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 32119  
LENGTH: 1114  
TYPE: DNA  
ORGANISM: Triticum aestivum  
US-10-953-349-32119

Query Match 5.3%; Score 30; DB 6; Length 1114;  
Best Local Similarity 53.4%; Pred. No. 1.5;  
Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 379 TGTGTTGTAACGACGCGGACGCACTCTTATGGAACAGAGCGCCCTTCATATCA 438  
DB 955 TGTGTTGTAACGACGAGACGAGACGAGCTGTCAAAAGCGCGGCGAGCGGC 896  
QY 439 GCGCGCGTTATCTATGCGCGGTGACCGGACAGAGGCGCGCTCCCTTATCGGC 496

D<sub>b</sub> 895 GCCGATGATTGCGTCTCGTCCGGGTCTGACGGGCGGGCTGCGCTCCCGCGC 838

```

RESULT 15
US-10-511-937-430/c
: Sequence 430. Application US/10511937
: Publication No. US20060088836A1
: GENERAL INFORMATION:
: APPLICANT: EXPRESSION DIAGNOSTICS, INC.
: APPLICANT: Wohlgenuth, Jay
: APPLICANT: Fry, Kirk
: APPLICANT: Woodward, Robert
: APPLICANT: Ly, Ngoc
: APPLICANT: Prentice, James
: APPLICANT: Morris, MacDonald
: APPLICANT: Rosenberg, Steven
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
: TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
: FILE REFERENCE: 506612000104
: CURRENT APPLICATION NUMBER: US/10/511,937
: CURRENT FILING DATE: 2004-10-19
: PRIOR APPLICATION NUMBER: PCT/US2003/012946
: PRIOR FILING DATE: 2003-04-24
: PRIOR APPLICATION NUMBER: US 10/131,831
: PRIOR FILING DATE: 2002-04-24
: PRIOR APPLICATION NUMBER: US 10/325,899
: PRIOR FILING DATE: 2002-12-20
: NUMBER OF SEQ ID NOS: 3117
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 430
: LENGTH: 1237
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-511-937-430

```

Query Match	5.3%;	Score 30;	DB 6;	Length 1237;
Best Local Similarity	67.7%;	Pred. No. 1.6;		
Matches 42;	Conservative	0;	Mismatches 20;	Indels 0;
			Gaps	0;

Gy		26	TAAATGGCTTCAACAACAAATTGCTTTGAAGTCGTTCATGTTGCCAACAGCACCCT	85
Dδ		1228	TAAATGGCTTGAAGAATCTTAATGGTAGAAGCCGATTTTTCATGTTTGCCCGTAAGCTCCT	11655

QY	86	TT	87
Db	1168	GT	1167

Search completed: June 15, 2006, 18:04:16  
Job time : 64.413 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 16:18:24 ; Search time 189,421 Seconds  
(without alignments)  
6515,052 Million cell updates/sec

Title: US-09-896-888a-1\_COPY\_351\_527

Perfect score: 177

Sequence: 1 gcttatacgcagacagcgc.....gccccaacgactctgtaaa 177

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_8:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*
- 14: geneseqn2005as:\*
- 15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177	100.0	560	12	ADQ48575
2	177	100.0	564	2	AAV62487
3	177	100.0	2773	9	AAU61306
4	175.4	99.1	5038	12	ADQ48539
5	41	23.2	141	12	ADQ48576
6	32	18.1	921	3	AAA02474
7	31	17.5	2109	14	ADY37449
8	30.8	17.4	1268	4	ABL18801
9	30.8	17.4	1342	4	ABL27503
10	30.8	17.4	3448	4	ABL27502
11	30.8	17.4	3600	4	ABL18800
12	30.6	17.3	29169	10	AAU64735
13	30.6	17.3	1404	13	ADT47065
14	30.6	17.3	10968	14	AECT5808
15	30.6	17.3	78869	14	AECT5774
16	30.2	17.1	1127	14	AEET6748
17	30	16.9	657	12	AD145632
18	30	16.9	1710	11	ABD14088

C 19	30	16.9	2485	15	AEP55630	Aef55630 Human gen
C 20	30	16.9	3107	13	ADS89385	Ad89385 Oligonucle
C 21	30	16.9	3501	13	ADS89483	Ad89483 Oligonucle
C 22	30	16.9	6107	6	ABK31430	Abk31430 Signal tr
C 23	30	16.9	6107	6	ABL70389	Ab170389 Chemical
C 24	30	16.9	6107	6	AA651341	AA651341 Human gen
C 25	30	16.9	132544	13	ADV99887	Adv99887 Nanchangm
C 26	29.8	16.8	1311	14	ACU65908	Ac165908 M. xanthu
C 27	29.8	16.8	1370	6	ADP83430	Adp83430 Bread whe
C 28	29.8	16.8	1789	4	ABL19245	Ab119245 Drosophi
C 29	29.8	16.8	3915	4	ABL19244	Ab119244 Drosophi
C 30	29.8	16.8	13579	14	ACL64638	ACL64638 M. xanthu
C 31	29.6	16.7	300	3	AAA01209	AAA01209 Human col
C 32	29.6	16.7	1798	13	ADT18768	Adt18768 Plant CDN
C 33	29.2	16.5	1593	13	ADT45665	Adt45665 Bacteri
C 34	28.6	16.2	585	11	ABD13766	Abd13766 Pseudom
C 35	28.6	16.2	823	12	ACH89781	Ach89781 Human gen
C 36	28.6	16.2	915	5	AA870780	AA870780 Human gen
C 37	28.6	16.2	1425	4	AA851490	AA851490 Pseudom
C 38	28.6	16.2	1425	8	ACA19456	ACA19456 Prokaryot
C 39	28.6	16.2	1434	13	ADS56860	Ad56860 Bacteri
C 40	28.6	16.2	1443	11	ABD13493	Abd13493 Pseudom
C 41	28.6	16.2	1551	11	ABD13719	Abd13719 Pseudom
C 42	28.6	16.2	2085	11	ABD13633	Abd13633 Pseudom
C 43	28.6	16.2	2088	15	AER86025	Aer86025 Human neu
C 44	28.6	16.2	2091	2	AA06989	Aa06989 Human neu
C 45	28.6	16.2	2091	6	ABK90730	Abk90730 CDNA enco

## ALIGNMENTS

RESULT 1	ADQ48575	standard; DNA; 560 BP.
ID	ADQ48575	
XX	ADQ48575	
AC	09-SEP-2004 (first entry)	
DT	09-SEP-2004 (first entry)	
XX	OpIE2 promoter DNA sequence.	
DE	viral vector; recombination site; recombinant virus;	
XX	KM replication-defective particle generation; gene expression inhibition;	
KW	gene therapy vector; ds; OpIE2 promoter.	
OS	Unidentified.	
XX	WO2004009768-A2.	
PN	29-JAN-2004.	
XX	18-JUL-2003; 2003WO-US022437.	
PF	18-JUL-2002; 2002US-0396335P.	
PR	26-JUL-2002; 2002US-0396335P.	
PR	15-NOV-2002; 2002US-0427231P.	
PR	24-MAR-2003; 2003US-0456496P.	
PR	03-JUN-2003; 2003US-0474940P.	
XX	(INV-) INVTROGEN CORP.	
PA	(BENN) BENNETT R P.	
PA	(WELC) WELCH P J.	
PA	(HARW) HARWOOD S.	
PA	(MADD) MADDEN K.	
PA	(FRIM) FRIMPONG K.	
PA	(FRAN) FRANKS K E.	
XX	Bennett RP, Welch PJ, Harwood S, Madden K, Frimping K, Franke KE;	
XX	WPI; 2004-132944/13.	
DR	Novel nucleic acid molecule comprising all or a portion of one or more	
XX		
PT		

PT viral genome and further comprising two or more recombination sites that  
 PT do not substantially recombine with each other, useful as gene therapy.  
 XX  
 PS Disclosure, Fig 16; 555pp; English.

XX The invention comprises a nucleic acid molecule consisting of all or a  
 CC portion of one or more viral/baculoviral genomes and further containing  
 CC at least two recombination sites that do not substantially recombine with  
 CC each other. One or more of the recombination sites is capable of  
 CC undergoing recombination with a compatible recombination site in the  
 CC presence of one or more proteins active in lambda recombination. The  
 CC nucleic acid of the invention replicates in prokaryotic and eukaryotic  
 CC cells. The nucleic acid of the invention is useful for constructing a  
 CC recombinant virus, generating replication-defective particles,  
 CC preventing/inhibiting the expression of one or more genes in an organism,  
 CC and are useful as gene therapy vectors. The nucleic acid of the invention  
 CC is also useful for producing and expressing fusion polypeptides. The  
 CC present DNA sequence represents the OpIE2 promoter that was used in the  
 CC exemplification of the invention.  
 XX

SO Sequence 560 BP; 144 A; 156 C; 116 G; 144 T; 0 U; 0 Other;

Query Match 100.0%; Score 177; DB 12; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-48;  
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTACGACGAGCGGAGCTTCTGTGTTGCTTAACCGGACCGGACGAACTCCT 60  
 Db 355 GTCTTATCGTACGACGAGCGGAGCTTCTGTGTTGCTTAACCGGACCGGACGAACTCCT 414

QY 61 TATCGAAGACGAGACGCGCCTTCATATGACCGCGGCTTATCTCATGCGCGTACCGGACA 120  
 Db 415 TATCGAAGACGAGACGCGCCTTCATATGACCGCGGCTTATCTCATGCGCGTACCGGACA 474

QY 121 CGAGGCGCCCGTCCGCTTATCGCGCTTAATAATACAGCCCGCAACGATCTGTTAA 177  
 Db 475 CGAGGCGCCCGTCCGCTTATCGCGCTTAATAATACAGCCCGCAACGATCTGTTAA 531

RESULT 2  
 AAV62487  
 ID AAV62487 standard; DNA; 564 BP.

XX AAV62487;  
 AC 17-OCT-2003 (revised)  
 DT 19-JAN-1999 (first entry)

XX O. pseudotsugata multicausid nucleopolyhedrosis virus 1e2 promoter.

XX Orygia pseudotsugata; multicausid; nucleopolyhedrosis virus; OpMNIV;  
 KM Op 1e2; promoter; shuttle vector; transformation; melanotransferrin;  
 KM immediate early baculovirus promoter; prokaryotic; transcription;  
 KW bleomycin/phenylethylamine-type antibiotic; insect cell; transposon;  
 KM ion transport peptide hormone; ss.

XX Orygia pseudotsugata; polyhedrosis virus.

OS WO9844141-A2.

PN 08-OCT-1998.

XX 26-MAR-1998; 98WO-CA000282.

XX 27-MAR-1997; 97US-0049946P.

XX 28-JAN-1998; 98CA-02221819.

XX (UYBR-) UNIV BRITISH COLUMBIA.

PI Grigoliacti TA, Theilmann DA, Pfeifer TA, Hegedus DD;  
 DR WPI, 1998-557129/47.

PT Expression vectors for transforming insect cells from disparate lines -  
 PT useful to express heterologous DNA, e.g. to allow study of gene  
 PT expression and produce commercially important proteins.  
 XX  
 XX Claim 10; Page 82; 121pp; English.

XX This represents a Orygia pseudotsugata multicausid nucleopolyhedrosis  
 CC virus (OpMNIV) immediate early 2 (Op 1e2) promoter sequence. The  
 CC invention provides a new shuttle vector for transforming insect cells  
 CC that comprises: (i) prokaryotic origin of replication; (ii) insect  
 CC promoter having homology to, and capable of functioning as, an immediate  
 CC early baculovirus promoter; (iii) prokaryotic promoter sequence, and (iv)  
 CC selectable marker capable of conferring resistance to a bleomycin/  
 CC phenylethylamine-type antibiotic under transcriptional control of (iii) and  
 CC (iii), in insect and prokaryotic cells respectively. The vectors can be  
 CC used to stably transform (especially insect) cells with heterologous DNA,  
 CC useful to allow study of gene expression and direct expression of  
 CC heterologous gene products e.g. commercially important proteins. They are  
 CC especially useful to allow expression of the heterologous  
 CC melanotransferrin, ion transport peptide hormones or biologically active  
 CC derivatives in insect cells. They enable transformation of insect cell  
 CC lines from disparate species, allowing screening of lines for optimum  
 CC post-translational modification of particular proteins. Shuttle vectors  
 CC further comprising DNA transposable elements defining a transposon can be  
 CC used to optimise heterologous protein expression and facilitate selection  
 CC of desired transformants. (Updated on 17-OCT-2003 to standardise OS  
 CC field)

SO Sequence 564 BP; 147 A; 160 C; 116 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 177; DB 2; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-48;  
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTACGACGAGCGGAGCTTCTGTGTTGCTTAACCGGACCGGACGAACTCCT 60  
 Db 351 GTCTTATCGTACGACGAGCGGAGCTTCTGTGTTGCTTAACCGGACCGGACGAACTCCT 410

QY 61 TATCGAAGACGAGACGCGCCTTCATATGACCGCGGCTTATCTCATGCGCGTACCGGACA 120  
 Db 411 TATCGAAGACGAGACGCGCCTTCATATGACCGCGGCTTATCTCATGCGCGTACCGGACA 470

QY 121 CGAGGCGCCCGTCCGCTTATCGCGCTTAATAATACAGCCCGCAACGATCTGTTAA 177  
 Db 471 CGAGGCGCCCGTCCGCTTATCGCGCTTAATAATACAGCCCGCAACGATCTGTTAA 527

RESULT 3  
 AAL61306  
 ID AAL61306 standard; DNA; 2773 BP.

XX AAL61306;

AC 22-SEP-2003 (first entry)

XX p2ZOp2F expression vector for insect cells.

XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;  
 KM tumour necrosis factor alpha; gene therapy; arthritis; gene; ds.

XX Unidentified.

PN WO2003042244-A2.

XX 22-MAY-2003.

XX 15-NOV-2002; 2002WO-DK000764.

XX 16-NOV-2001; 2001DK-00001702.

XX 16-NOV-2001; 2001US-0331575P.

PA (PHAR-) PHARMEXA AS.  
 PA (KLVS/) KLYSNER S.



PA (NIEL/) NIELSEN F. S.  
PA (BRAT/) BRATT T.  
PA (VOLD/) VOLDORGB. B.  
PA (MOUR/) MOURITSEN S.  
PI Klynsner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;  
DR WPI; 2003-449558/42.  
XX  
XX  
PT New immunogenic analogue of a polymeric protein, useful for preparing a  
composition for treating inflammatory diseases e.g. arthritis.  
XX  
XX PS Disclosure; Page 195-196; 196pp; English.  
XX  
XX The invention relates to immunogenic analogues of multimeric proteins  
such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis  
factor alpha (TNF, TNFalpha) and methods for production of immunogenic  
analogues. The immunogenic analogue is useful for preparing a composition  
for treating inflammatory diseases, e.g., arthritis. It is also used in  
gene therapy. The present sequence is p2ZOP2P expression vector for  
CC insect cells. This sequence is used to illustrate the method of the  
invention  
CC  
XX  
SQ Sequence 2773 BP; 669 A; 750 C; 673 G; 681 T; 0 U; 0 Other;  
Query Match 100.0%; Score 177; DB 9; Length 2773;  
Best Local Similarity 100.0%; Pred. No. 4.5e-48;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTCCTATCGTGACGAGCGCGCCTCTGTTGCTTACCGGACGGGCAACTCCT 60  
DB 355 GTCTTATCGTGACGAGCGCGCCTCTGTTGCTTACCGGACGGGCAACTCCT 414  
QY 61 TATCGGAACGAGCGCGCCTCATATCAGCCGCGGTTATCTCATGCGCGTACCGGACA 120  
DB 415 TATCGGAACGAGCGCGCCTCATATCAGCCGCGGTTATCTCATGCGCGTACCGGACA 474  
QY 121 CGAGGCGCGCGTCCGCTTATCGGCGCTTAAATACGCCCGGCAAGATCTGTTAA 177  
DB 475 CGAGGCGCGCGTCCGCTTATCGGCGCTTAAATACGCCCGGCAAGATCTGTTAA 531  
RESULT 4  
ADQ48539  
ID ADQ48539 standard; DNA; 5038 BP.  
XX  
AC ADQ48539;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Viral vector-related plasmid - PIB/V5-His-DEST.  
XX  
XX viral vector; recombination site; recombinant virus;  
KW replication-defective particle generation; gene expression inhibition;  
KM gene therapy vector; ds; plasmid.  
XX  
OS Unidentified.  
XX  
XX WO2004009768-A2.  
XX  
PD 29-JAN-2004.  
XX  
PF 18-JUL-2003; 2003WO-US022437.  
XX  
XX 18-JUL-2002; 2002US-0396335P.  
PR 26-JUL-2002; 2002US-0398617P.  
PR 19-NOV-2002; 2002US-0427231P.  
PR 24-MAR-2003; 2003US-0456496P.  
PR 03-JUN-2003; 2003US-0474940P.  
XX  
XX (INVT-) INVITROGEN CORP.  
PA (BENN/) BENNETT R P.  
PA (WELC/) WELCH P J.

PA (HAR/) HARWOOD S.  
PA (MADD/) MADDEN K.  
PA (FRIM/) FRIMPONG K.  
PA (FRAN/) FRANK K E.  
PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;  
DR WPI; 2004-132944/13.  
XX  
XX  
XX Novel nucleic acid molecule comprising all or a portion of one or more  
PT viral genome and further comprising two or more recombination sites that  
PT do not substantially recombine with each other, useful as gene therapy.  
XX  
XX PS Example 18; Page 395-403; 555pp; English.  
XX  
XX The invention comprises a nucleic acid molecule consisting of all or a  
portion of one or more viral/baculoviral genomes and further containing  
CC at least two recombination sites that do not substantially recombine with  
CC each other. One or more of the recombination sites is capable of  
CC undergoing recombination with a compatible recombination site in the  
CC presence of one or more proteins active in lambda recombination. The  
CC nucleic acid of the invention replicates in prokaryotic and eukaryotic  
CC cells. The nucleic acid of the invention is useful for constructing a  
CC recombinant virus, generating replication-defective particles,  
CC preventing/inhibiting the expression of one or more genes in an organism,  
CC and are useful as gene therapy vectors. The nucleic acid of the invention  
CC is also useful for producing and expressing fusion polypeptides. The  
CC present DNA sequence represents a plasmid that was used in the  
exemplification of the invention.  
XX  
SQ Sequence 5038 BP; 1342 A; 1178 C; 1198 G; 1320 T; 0 U; 0 Other;  
Query Match 99.1%; Score 175.4; DB 12; Length 5038;  
Best Local Similarity 99.4%; Pred. No. 1.8e-47;  
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GTCCTATCGTGACGAGCGCGCCTCTGTTGCTTACCGGACGGGCAACTCCT 60  
DB 351 GTCTTATCGTGACGAGCGCGCCTCTGTTGCTTACCGGACGGGCAACTCCT 410  
QY 61 TATCGGAACGAGCGCGCCTCATATCAGCCGCGGTTATCTCATGCGCGTACCGGACA 120  
DB 411 TATCGGAACGAGCGCGCCTCATATCAGCCGCGGTTATCTCATGCGCGTACCGGACA 470  
QY 121 CGAGGCGCGCGTCCGCTTATCGGCGCTTAAATACGCCCGGCAAGATCTGTTAA 177  
DB 471 CGAGGCGCGCGTCCGCTTATCGGCGCTTAAATACGCCCGGCAAGATCTGTTAA 527  
RESULT 5  
ADQ48576  
ID ADQ48576 standard; DNA; 141 BP.  
XX  
AC ADQ48576;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Viral vector-related plasmid PIB/V5-His-DEST recombination region #1.  
XX  
XX viral vector; recombination site; recombinant virus;  
KW replication-defective particle generation; gene expression inhibition;  
KM gene therapy vector; ds; plasmid; recombination region.  
XX  
OS Unidentified.  
XX  
XX WO2004009768-A2.  
XX  
PD 29-JAN-2004.  
XX  
PF 18-JUL-2003; 2003WO-US022437.  
XX  
XX 18-JUL-2002; 2002US-0396335P.  
PR 26-JUL-2002; 2002US-0398617P.

```
PR 19-NOV-2002; 2002US-0427231P.  
PR 24-MAR-2003; 2003US-045496P.  
PR 03-JUN-2003; 2003US-0474940P.  
XX  
XX  
PA (INVT-) INVITROGEN CORP.  
PA (BENNT-) BENNETT R P.  
PA (WELC/) WELCH P J.  
PA (HARW/) HARWOOD S.  
PA (MADD/) MADDEN K.  
PA (FRIM/) FRIMPONG K.  
PA (FRAN/) FRANK K E.  
XX  
XX  
PI Bennett RP, Welch PJ, Harwood S, Madden K, Frimpong K, Franke KE;  
XX  
XX WPI; 2004-132944/13.  
XX  
XX Novel genome acid molecule comprising all or a portion of one or more  
PT viral genome and further comprising two or more recombination sites that  
PT do not substantially recombine with each other, useful as gene therapy.  
XX  
XX  
PS Disclosure; Fig 17; 555pp; English.  
XX  
XX The invention comprises a nucleic acid molecule consisting of all or a  
CC portion of one or more viral/baculoviral genomes and further containing  
CC at least two recombination sites that do not substantially recombine with  
CC each other. One or more of the recombination sites is capable of  
CC undergoing recombination with a compatible recombination site in the  
CC presence of one or more proteins active in lambda recombination. The  
CC nucleic acid of the invention replicates in prokaryotic and eukaryotic  
CC cells. The nucleic acid of the invention is useful for constructing a  
CC recombinant virus, generating replication-defective particles,  
CC preventing/inhibiting the expression of one or more genes in an organism,  
CC and are useful as gene therapy vectors. The nucleic acid of the invention  
CC is also useful for producing and expressing fusion polypeptides. The  
CC present DNA sequence represents a recombination region of a plasmid that  
CC was used in the exemplification of the invention.  
XX  
XX  
SQ Sequence 141 BP; 43 A; 36 C; 27 G; 34 T; 0 U; 1 Other;  
XX  
XX  
Query Match 23.2%; Score 41; DB 12; Length 141;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 137 CTTATCGCGCTATTAATACGCCGACAGATCTGTGTA 177  
Db 1 CTTATCGCGCTATTAATACGCCGACAGATCTGTGTA 41  
XX  
XX  
RESULT 6  
ID AAA02474/c  
XX AAA02474 standard; cDNA; 921 BP.  
XX  
XX AAA02474;  
XX  
XX 19-MAY-2000 (first entry)  
XX  
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2465.  
XX  
XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;  
XX detection; cancerous state; metastasis; identification; breast cancer;  
XX oestrogen receptor-positive breast cancer; therapy;  
XX oestrogen receptor-negative breast cancer; lung cancer; ss.  
XX  
XX Homo sapiens.  
XX  
XX PN WO958675-A2.  
XX  
XX PD 18-NOV-1999.  
XX  
XX PF 13-MAY-1999; 99MO-US010602.  
XX  
XX PR 14-MAY-1998; 98US-0085426P.  
XX PR 15-MAY-1998; 98US-0085537P.
```

```
PR 15-MAY-1998; 98US-0085696P.  
PR 21-OCT-1998; 98US-0105234P.  
PR 27-OCT-1998; 98US-0105877P.  
XX  
XX  
PA (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEO INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
PI Lamson G, Drmanac R, Ctkvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Steacie-Crain B;  
XX  
XX WPI; 2000-126369/11.  
XX  
XX Polynucleotide library used to determine cancerous states of mammalian  
PT cells.  
XX  
XX Claim 1; Page 989; 1097pp; English.  
XX  
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
CC libraries constructed from human colon cancer cell lines. The present  
CC invention also describes a method of detecting differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell, comprising  
CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was derived. The  
CC polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of pre-  
CC metastatic or metastatic cancerous states, stages of cancer or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer  
XX  
XX  
SQ Sequence 921 BP; 239 A; 214 C; 205 G; 223 T; 0 U; 40 Other;  
XX  
XX  
Query Match 18.1%; Score 32; DB 3; Length 921;  
Best Local Similarity 51.5%; Pred. No. 2.5;  
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
XX  
XX 16 GAGCCAGAGCTTCTGTGTGCTTAACCGACGCCGAGCGCAACTCTTATCGGAACAGAGAG 75  
Db 157 GGCAGCTGCGCNCACAGTGGCTCTCCACGCCGGTCTCCGGCTCCCGGCGGGT 98  
XX  
XX 76 CGCCTCCATATACGCCGCGGTATCTCATGCGCGTACCGAGACAGAGCGCCGTCCTCC 135  
Db 97 GCGGTCTCMAANCCACCTTCCTTGCTCTCCGCGCTCCGGTTCAGGAGCGACTGTCTT 38  
XX  
XX 136 GCTTATCGCGCC 147  
XX  
XX 37 TTAAGATCGNGCC 26  
XX  
XX  
RESULT 7  
ID ADJ37449/c  
XX ADJ37449 standard; DNA; 2109 BP.  
XX  
XX ADJ37449;  
XX  
XX 19-MAY-2005 (first entry)  
XX  
XX DNA encoding acyl-CoA oxidase.  
XX  
XX de; gene; diagnostic; acyl-CoA oxidase.  
XX  
XX OS Archaeobacter ureafaciens; IF012140.  
XX  
XX FH Key Location/Qualifiers  
XX CDS 1..2109  
XX FT /*tag= a
```

FT /product= "Acyl-CoA oxidase"  
 FT /partial  
 FT /note= "No Stop codon"  
 XX  
 XX WO2005019458-A1.  
 XX  
 XX 03-MAR-2005.  
 XX  
 XX 10-AUG-2004; 2004WO-JP011461.  
 XX  
 XX 25-AUG-2003; 2003JP-00299607.  
 XX  
 XX (KIKK ) KIKKOMAN CORP.  
 XX  
 XX Bakke M, Kajiyama N;  
 XX  
 XX WPI; 2005-214252/22.  
 XX  
 XX P-PSDB; ADY37448.  
 XX  
 XX Novel acyl-CoA oxidase exhibiting high stability in presence of thiol  
 PT reagent at specific pH and temperature, useful in clinical diagnosis.  
 XX  
 XX Claim 5; SEQ ID NO 2; 21pp; Japanese.  
 XX  
 XX The invention relates to Arthrobacter ureafaciens (IFO12140) acyl-CoA  
 CC oxidase (I), which has excellent stability in the presence of thiol  
 CC reagent at 37 degrees C. Also described are: the acyl-CoA oxidase gene  
 CC (II) encoding (I); recombinant DNA (III) obtained by introducing (II)  
 CC into a vector DNA; a transformed host or transduced host (IV) containing  
 CC (III); and preparation of (I). (I) is useful in clinical diagnosis. The  
 CC present sequence represents DNA encoding acyl-CoA oxidase. Note: The  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences/03.03.2005/.

XX  
 SQ Sequence 2109 BP; 401 A; 695 C; 662 G; 351 T; 0 U; 0 Other;  
 XX  
 XX Query Match 17.5%; Score 31; DB 14; Length 2109;  
 XX Best Local Similarity 52.8%; Pred. No. 6.4;  
 XX Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
 XX  
 QY 8 CGTGACGAGACGCGCTTCCTGTGCTAACCGGACGCGGCAACTCCTTATGGA 67  
 Db 1503 CGGACCTCGCGGAGGCTGTCCGCTGCAAGGCCACCCCGGACCTGCTCAGAGCTA 1444  
 QY 68 ACAGGACGCGCTTCATATACCGCGCGTTATTCATGCGCGTACCGGACGAGGCG 127  
 Db 1443 CCGGCGCAGGACGCGCAAGTTGGCGCCACGGAATCCTTGCGTATCGGCGAGGCG 1384  
 QY 128 CCGGTCC 134  
 Db 1383 CTTGGCC 1377

XX  
 XX RESULT 8  
 XX ABL18801  
 XX ID ABL18801 standard; DNA, 1268 BP.  
 XX  
 XX ABL18801;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7876.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX  
 XX WO200171042-A2.  
 PN  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 XX WPI; 2001-656860/75.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 XX Claim 1; SEQ ID NO 7876; 21pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
 CC AB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 1268 BP; 295 A; 336 C; 360 G; 277 T; 0 U; 0 Other;  
 XX  
 XX Query Match 17.4%; Score 30.8; DB 4; Length 1268;  
 XX Best Local Similarity 55.7%; Pred. No. 6.6;  
 XX Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
 XX  
 QY 17 ACGCAGCTTCCTGTGCTTAACCGGACGCGGCAACTCCTTATCGGACGAGCG 76  
 Db 657 ACAAGCTTCGCGTGGAGGTGCCACAGCGGCTCTAGTCTTGCGGCTCGGTGCG 716  
 QY 77 GCCTCATATCAGCGCGGCTTATTCATGCGGCTACCGGACG 122  
 Db 717 GTTGGACATTTGGCCATACCAATCTCAAGTCCAGAAAGTACAAG 762

XX  
 XX RESULT 9  
 XX ABL27503  
 XX ID ABL27503 standard; DNA; 1342 BP.  
 XX  
 XX ABL27503;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 33982.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX  
 XX WO200171042-A2.  
 PN  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 XX WPI; 2001-656860/75.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 33982; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1342 BP; 324 A; 347 C; 372 G; 299 T; 0 U; 0 Other;  
Query Match 17.4%; Score 30.8; DB 4; Length 1342;  
Best Local Similarity 55.7%; Pred. No. 6.7;  
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
QY 17 ACGCCAGCTTCTGTGTGCTAACCGGACCGGACGCACTCTTATCGGAACAGACGC 76  
DB 731 ACAACAGCTTCCGTTGAGGTGCGCCACAGCGGGTCTTATGCGGCTCCGTTGCG 790  
QY 77 GCCTCCATATCAGCGCGCGCTTATCTCATGCGCGTACCGGACG 122  
DB 791 GTTGGACATTTGGCCATACCAATCTCAAGTCCAGAAAGTACAG 836  
RESULT 10  
ABL27502/c  
ID ABL27502 standard; DNA; 3448 BP.  
XX  
AC ABL27502;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 33979.  
XX  
XX Drosophila; developmental biology; cell signaling; insecticide;  
XX pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001MO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEXE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 33979; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3448 BP; 894 A; 817 C; 801 G; 936 T; 0 U; 0 Other;  
Query Match 17.4%; Score 30.8; DB 4; Length 3448;  
Best Local Similarity 55.7%; Pred. No. 8.4;  
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
QY 17 ACGCCAGCTTCTGTGTGCTAACCGGACCGGACGCACTCTTATCGGAACAGACGC 76  
DB 1667 ACAACAGCTTCCGTTGAGGTGCGCCACAGCGGGTCTTATGCGGCTCCGTTGCG 1608  
QY 77 GCCTCCATATCAGCGCGCGCTTATCTCATGCGCGTACCGGACG 122  
DB 1607 GTTGGACATTTGGCCATACCAATCTCAAGTCCAGAAAGTACAG 1562  
RESULT 11  
ABL1800/c  
ID ABL1800 standard; DNA; 3600 BP.  
XX  
AC ABL1800;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7873.  
XX  
XX Drosophila; developmental biology; cell signaling; insecticide;  
XX pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001MO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX  
XX 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEXE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
XX Claim 1; SEQ ID NO 7873; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
CC AB572072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3600 BP; 946 A; 856 C; 823 G; 975 T; 0 U; 0 Other;  
Query Match 17.4%; Score 30.8; DB 4; Length 3600;  
Best Local Similarity 55.7%; Pred. No. 8.5;  
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 17 ACCGACCTTCCTGTTCTTACCGACCGGACGACCTCTTATCGAAGACGACG 76  
DB 1667 ACAACACCTTCCTGTTGAGGAGGCGCCACAGCGGGTCTTGTGCGCGCTCCGGTGGC 1608  
QY 77 GCCTCCATATCAGCGCGCGCTTATCTCATGCGCGTGAACCGACAGC 122  
DB 1607 GTTGGCACATTGGCCATACAAATCTCTCAAGTCCAGAAAGTACAG 1562

RESULT 12  
AAD64735  
ID AAD64735 standard; DNA; 29169 BP.  
XX  
AC AAD64735;  
XX  
XX 11-MAR-2004 (first entry)  
XX  
XX Human carcinoma-associated (CA) gene TBX21.  
XX  
XX  
XX Carcinoma-associated protein; CAP; acinar cell carcinoma; fibrosarcoma;  
KM Kapost's sarcoma; breast cancer; Hairy cell leukaemia; human; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US2003099963-A1.  
XX  
XX 29-MAY-2003.  
XX  
XX 20-MAR-2002; 2002US-00105613.  
XX  
XX 22-DEC-2000; 2000US-00747377.  
XX 02-MAR-2001; 2001US-00798586.  
XX 08-NOV-2001; 2001US-00052482.  
XX  
XX (MORR/) MORRIS D W.  
XX (ENGEL/) ENGELHARD E K.  
XX  
XX Morris DW, Engelhard EK;  
XX  
XX WPI; 2003-874605/81.  
XX  
XX Novel recombinant carcinoma-associated protein such as mouse or human  
PT TBX21 protein, useful for screening a bioactive agent capable of binding  
PT to carcinoma associated protein, and for evaluating the effect of a  
PT candidate carcinoma drug.  
XX  
XX  
XX Disclosure; SEQ ID NO 4; 0pp; English.  
XX  
XX The present invention relates to novel recombinant carcinoma-associated  
XX protein (CAP) useful for screening a bioactive agent capable of binding  
XX to carcinoma associated (CA) protein and for evaluating the effect of a  
XX candidate carcinoma drug. The invention also describes the use of novel  
XX compositions for use in screening methods and provides compositions and  
XX methods associated with altered expression of TBX21 in cancer. Suitable  
XX cancers which can be diagnosed or screened by the invention includes  
XX acinar cell carcinoma, fibrosarcoma, Kapost's sarcoma, breast cancer and  
XX Hairy cell leukaemia. The present sequence is human carcinoma-associated  
XX (CA) gene  
XX  
SQ Sequence 29169 BP; 5725 A; 7553 C; 7224 G; 8641 T; 0 U; 26 Other;

Query Match 17.4%; Score 30.8; DB 10; Length 29169;  
Best Local Similarity 61.0%; Pred. No. 14;  
Matches 50; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 11 GACAGACGCGCAGCTTCCTGTTGCTAACCGACCGGAGCGCAACTCTTATCGAACA 70  
DB 8982 GACACGACGCGCTTCCTGTTGCTAACCGACCGGAGCGCAACTCTTATCGGAG 9041  
QY 71 GACGCGCGCTTCATATCAGCGC 92  
DB 9042 CCCCCTGCGCCACCTCCCGC 9063

RESULT 13  
ADT47065  
ID ADT47065 standard; cDNA; 1404 BP.  
XX  
XX ADT47065;  
XX  
XX 02-DEC-2004 (first entry)  
XX  
XX  
XX Bacterial polynucleotide #21816.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
KM pathogen tolerance; pest tolerance; plant disease resistance;  
KM cell cycle pathway modification; plant growth regulator;  
KM homologous recombination; seed oil yield; protein yield; carbohydrate;  
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KM bacterial polynucleotide; gene; ss.  
XX  
XX Bacteria.  
OS  
XX  
XX US2003233675-A1.  
XX  
XX 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 45503; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source. The invention also relates to a transformed plant  
XX comprising the recombinant DNA construct and a method of producing a  
XX transformed plant having an improved property. The plant is a crop plant  
XX such as maize or soybean. The method of producing a transformed plant  
XX having an improved property comprises transforming a plant with the  
XX recombinant DNA construct and growing the transformed plant, where the  
XX polynucleotide or polypeptide is useful for improving plant properties.  
XX The recombinant DNA construct is useful for producing plants with  
XX improved plant properties, e.g. improved cold, heat or drought tolerance,  
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
XX increased resistance to plant disease, better growth rate by modification  
XX of the cell cycle pathway with plant growth regulators, increased rate of  
XX homologous recombination, modified seed oil or protein yield and/or  
XX phosphorus use and/or uptake, by modification of photosynthesis or by  
XX providing improved plant growth and development under at least one stress  
XX condition, improved lignin production or improved galactomannan  
XX production. This sequence represents a bacterial polynucleotide used in  
XX the scope of the invention. Note: The sequence data for this patent did  
XX not form part of the printed specification but was obtained in electronic  
XX format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 1404 BP; 236 A; 489 C; 454 G; 225 T; 0 U; 0 Other;

```

Query Match          17.3%; Score 30.6; DB 13; Length 1404;
Best Local Similarity 62.3%; Pred. No. 7.9;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY      99 ATCTCATGCGGTGACCGGACGAGCGCCGTCCTTATGCGGCTTAATAATACAG 158
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      464 ACCGACGAGGGGTGAACGTGCCGATGTGTGTCGCGCTTGCGGCTTTGCGCAAG 523
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      159 CCGGCAACGATCTGTGA 175
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      524 ACCGCGACGATCTGGAA 540
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
ID      AEC75808 standard; DNA; 10968 BP.
AC      AEC75808;
XX
XX
DT      01-DEC-2005 (first entry)
XX
XX      P. celluloseum ambruticin ambC gene.
XX
XX      de; gene; ambC; ambruticin; fungal infection; fungicide; polyketide.
XX
XX      Polyangium celluloseum.
XX
XX      WO2005086907-A2.
XX
XX      22-SEP-2005.
XX
XX      08-MAR-2005; 2005WO-US007924.
XX
XX      08-MAR-2004; 2004US-0551103P.
XX      PR      04-MAY-2004; 2004US-0568290P.
XX      PR      07-MAR-2005; 2005US-00075185.
XX
XX      (KOSA-) KOSAN BIOSCIENCES INC.
XX
XX      Reeves CD, Julien B, Reid RC;
XX
XX      WPI; 2005-639088/65.
XX
XX
XX      New nucleic acids encoding polyketide synthases and polyketide
XX      PT      modification enzymes involved in or required for the biosynthesis of
XX      PT      ambruticin, useful for producing polyketides, e.g. ambruticins, for
XX      PT      treating fungal infections.
XX
XX      Claim 4; SEQ ID NO 35; 2855bp; English.
XX
XX
XX      The invention relates to a purified or recombinant nucleic acid (I)
XX      CC      comprising a nucleotide sequence encoding a polypeptide involved in or
XX      CC      required for the biosynthesis of ambruticin. The complement of the
XX      CC      nucleotide sequence hybridizes to a nucleic acid encoding a domain
XX      CC      selected from beta-ketoacyl synthase, acyltransferase, beta-ketoreductase,
XX      CC      dehydratase, enoylreductase, or acyl carrier protein domains of the
XX      CC      Sorangium cellulosum (synonym: Polyangium celluloseum) ambruticin gene
XX      CC      cluster, or where the complement of the nucleotide sequence hybridizes to
XX      CC      a nucleic acid encoding an amino acid sequence selected from (SEQ ID NO.
XX      CC      11-19 (AEC75784-AEC75792). The nucleic acid is useful for producing
XX      CC      polyketides, including ambruticins and ambruticin analogs and
XX      CC      derivatives. Ambruticins can be used for treating fungal infections. The
XX      CC      present sequence represents P. celluloseum ambruticin ambC gene.
XX
XX      Sequence 10968 BP; 1231 A; 3980 C; 4394 G; 1363 T; 0 U; 0 Other;
SQ

Query Match          17.3%; Score 30.6; DB 14; Length 10968;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 88; Conservative 0; Mismatches 74; Indels 3; Gaps 1;

QY      3 CTATATGTAACAGAGCCGACCTTCTGTGTGTTGCTAACCGGACGCGGACCAACTCTTTA 62
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      6775 CTCTCTCTGTGGGTGCGGACGAGCCGCGCTGCGGCCGACGCGGCGGCTCGCC--- 6831
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY      63 TCGGAACAGACGCGCTCCATATCAGCGCGGCTTATCTCAGCGGTGACCGGACAG 122
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      6832 GAGACACTGCGCGCCACCCGACAGCGGCTCTCTGACCTGCCCGGACCTGGCCACG 6891
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      123 AGCGCGCCGTCCTTATGCGGCTTAATAATACAGCCGCAAG 167
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      6892 ACCGCGACGACCTCGGCAACGCGGCTCGCGCTCCGCTCGCGCG 6936
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
ID      AEC75774 standard; DNA; 78869 BP.
AC      AEC75774;
XX
XX
DT      01-DEC-2005 (first entry)
XX
XX      Polyangium celluloseum ambruticin biosynthetic gene cluster.
XX
XX      de; ambruticin; fungal infection; fungicide; polyketide.
XX
XX      Polyangium celluloseum.
XX
XX      WO2005086907-A2.
XX
XX      22-SEP-2005.
XX
XX      08-MAR-2005; 2005WO-US007924.
XX
XX      08-MAR-2004; 2004US-0551103P.
XX      PR      04-MAY-2004; 2004US-0568290P.
XX      PR      07-MAR-2005; 2005US-00075185.
XX
XX      (KOSA-) KOSAN BIOSCIENCES INC.
XX
XX      Reeves CD, Julien B, Reid RC;
XX
XX      WPI; 2005-639088/65.
XX
XX      P-PSDB; AEC75776, AEC75777, AEC75778, AEC75779, AEC75780, AEC75781,
XX      DR      AEC75782, AEC75783, AEC75784, AEC75785, AEC75786, AEC75787, AEC75788,
XX      DR      AEC75789, AEC75790, AEC75791, AEC75792, AEC75801, AEC75802, AEC75803.
XX
XX      New nucleic acids encoding polyketide synthases and polyketide
XX      PT      modification enzymes involved in or required for the biosynthesis of
XX      PT      ambruticin, useful for producing polyketides, e.g. ambruticins, for
XX      PT      treating fungal infections.
XX
XX      Disclosure; SEQ ID NO 1; 2855bp; English.
XX
XX
XX      The invention relates to a purified or recombinant nucleic acid (I)
XX      CC      comprising a nucleotide sequence encoding a polypeptide involved in or
XX      CC      required for the biosynthesis of ambruticin. The complement of the
XX      CC      nucleotide sequence hybridizes to a nucleic acid encoding a domain
XX      CC      selected from beta-ketoacyl synthase, acyltransferase, beta-ketoreductase,
XX      CC      dehydratase, enoylreductase, or acyl carrier protein domains of the
XX      CC      Sorangium cellulosum (synonym: Polyangium celluloseum) ambruticin gene
XX      CC      cluster, or where the complement of the nucleotide sequence hybridizes to
XX      CC      a nucleic acid encoding an amino acid sequence selected from (SEQ ID NO.
XX      CC      11-19 (AEC75784-AEC75792). The nucleic acid is useful for producing
XX      CC      polyketides, including ambruticins and ambruticin analogs and
XX      CC      derivatives. Ambruticins can be used for treating fungal infections. The
XX      CC      present sequence represents Polyangium celluloseum ambruticin biosynthetic
XX      CC      gene cluster.
XX
XX      Sequence 78869 BP; 10766 A; 27557 C; 28921 G; 11625 T; 0 U; 0 Other;
SQ

Query Match          17.3%; Score 30.6; DB 14; Length 78869;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 88; Conservative 0; Mismatches 74; Indels 3; Gaps 1;

QY      3 CTATATGTAACAGAGCCGACCTTCTGTGTGTTGCTAACCGGACGCGGACCAACTCTTTA 62
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 22042 CTGCTCTGTGTGGGTTCGGACGAGCCGGCCGCTGGCCGCCAGGCCGGCGGCTCGCC--- 22098  
OY 63 TCGGAACAGAGACGGGCTCTCCATATCAGCCGGCGCTTATCTCATGCGCGGTGACCGGACACG 122  
Db 22099 GAGCACCCTGCGCGGCCCAACCCGAGCCAGGGGCTGCTCGACGTGCGCCGCGAGCCTGGCCACG 22158  
OY 123 AGGGGCCCGTCCGGCTTATGCGGCTATAAATACAGCCCGCACAG 167  
Db 22159 ACGCGACGACACTTCGCCACGGGGCTGCGCTGCGCTCGGCGCG 22203

Search completed: June 15, 2006, 16:31:52  
Job time : 192.921 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 16:32:06 ; Search time 1481.21 Seconds  
(without alignments)  
6682.184 Million cell updates/sec

Title: US-09-896-888a-1\_COPY\_351\_527

Perfect score: 177

Sequence: 1 gctctatcgagacagcgc.....gcccgcaacgactctgtraa 177

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc3:\*  
3: gb\_esc4:\*  
4: gb\_esc5:\*  
5: gb\_esc6:\*  
6: gb\_esc7:\*  
7: gb\_esc2:\*  
8: gb\_esc8:\*  
9: gb\_esc9:\*  
10: gb\_esc10:\*  
11: gb\_gsa1:\*  
12: gb\_gsa2:\*  
13: gb\_gsa3:\*  
14: gb\_gsa4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.4	20.0	822	9	DN873866 nad29c08.
2	33.8	19.1	507	11	AQ297217 HS_2266_A
3	33.6	19.0	343	4	BY342594 BY342594
4	33.6	19.0	402	4	BY077721 BY077721
5	33.2	18.8	931	8	CV068875 f2_new.ch
6	33	18.6	348	4	BY338426 BY338426
7	33	18.6	353	4	BY337247 BY337247
8	33	18.6	354	4	BY013584 BY013584
9	33	18.6	355	4	BY340467 BY340467
10	33	18.6	360	4	BY343718 BY343718
11	33	18.6	362	4	BY339252 BY339252
12	33	18.6	371	4	BY083658 BY083658
13	33	18.6	611	5	CP916169 B0991A12-
14	33	18.6	621	5	B0620008 TAr1138G-
15	33	18.6	624	5	CP913548 B0950D02-
16	33	18.6	635	7	AA989838 uc79f07.y
17	33	18.6	641	1	AA989838 uc79f07.y
18	33	18.6	657	5	CP915946 B0987F04-
19	33	18.6	669	5	CP174622 B0941F09-

20	33	18.6	973	4	CA490927	CA490927
21	33	18.6	1029	13	CA919246	CA919246
22	33	18.6	1987	6	AK144659	AK144659
23	32.8	18.5	617	13	CM192940	CM192940
24	32.8	18.5	638	12	CE066545	CE066545
25	32.6	18.4	786	13	CL421872	CL421872
26	32.4	18.3	636	8	CO973145	CO973145
27	32	18.1	327	4	BY789720	BY789720
28	32	18.1	1297	3	BQ645920	BQ645920
29	31.8	18.0	553	8	CR540072	CR540072
30	31.8	18.0	1076	6	BA463482	BA463482
31	31.8	18.0	3592	6	CR859367	CR859367
32	31.6	17.9	681	14	DX006812	DX006812
33	31.6	17.9	1011	13	CL494562	CL494562
34	31.6	17.9	1310	3	BQ720271	BQ720271
35	31.4	17.7	344	4	BY340052	BY340052
36	31.4	17.7	361	5	CJ079027	CJ079027
37	31.4	17.7	374	7	BB843476	BB843476
38	31.4	17.7	392	5	CJ086368	CJ086368
39	31.4	17.7	681	12	CG295117	CG295117
40	31.2	17.6	543	5	CK071906	CK071906
41	31.2	17.6	595	5	CK066979	CK066979
42	31.2	17.6	665	13	CM170205	CM170205
43	31.2	17.6	666	13	CM502443	CM502443
44	31.2	17.6	693	12	CL190317	CL190317
45	31.2	17.6	787	13	CM028427	CM028427

## ALIGNMENTS

RESULT 1  
DN873866 822 bp mRNA linear EST 21-Apr-2005  
nad29c08.y1 Dog eye cornea. Unnormalized (nad) Canis familiaris  
LOCUS  
DEFINITION  
CDNA clone nad29c08 5', mRNA sequence.  
ACCESSION  
DN873866  
VERSION  
DN873866.1 GI:62843807  
KEYWORDS  
EST.  
SOURCE  
Canis familiaris (dog)  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
Canis.

REFERENCE  
1. (bases 1 to 822)  
Wistow, G.  
NEIBank analysis of Dog cornea  
Unpublished (2005)  
JOURNAL  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 29 row: c column: 08  
Seq primer: Universal M13 Reverse.  
Location/Qualifiers  
1. 822

FEATURES  
source  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/strain="Beagle"  
/db\_xref="taxon:9615"  
/clone="nad29c08"  
/issue\_type="Cornea"  
/dev\_stage="Adult"  
/lab\_pos="EMDH108"  
/clone\_lib="Dog eye cornea. Unnormalized (nad)"  
/note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from dog cornea tissue. A directionally cloned cDNA library in the pCMVSPORT6 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript

Plasmid System, full details of which are contained in the manufacturer's instruction manual (<http://www.lifetechn.com/>). First strand synthesis was carried out using a Not I primer-adaptor (5'-pGACTAGTCTAGATCGCGAGCGCGCC(T)15-3'). cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC). Analyzed data available through <http://neibank.nhl.nih.gov>.

## ORIGIN

Query Match 20.0%; Score 35.4; DB 9; Length 822;

Best Local Similarity 54.1%; Pred. No. 2.4;

Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 13 CAGAGCCGACGCTTCTGTGCTTACCGGACCGGACGCTTATCGGAACAG 72  
DB 81 CAGGCGCCGACGATGATGTCGCGGACGCGGCGGCTCCGCGCGCGCG 140  
QY 73 ACGGCGCTCCATATCAGCGCGGCTTATCTCATGCGGTACCGGACGAGCGCGCGT 132  
DB 141 CTTCGCGGACTTCTTCGCTCGAGCTCTCCAGCGCGCGGATCTCTGAGCGGTCCT 200  
QY 133 CCGCGTTATCGCG 145  
DB 201 CCAGATTCGCGG 213

## RESULT 2

LOCUS

AC297217 507 bp DNA linear GSS 15-DEC-1998  
HS 2266 A1 A08 MR CIT Approved Human Genomic Sperm Library D Homo

## DEFINITION

Sequence genomic clone Plate=2266 Col=15 Row=A, genomic survey

## ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 507)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

10449764

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2266 row: A column: 15

Class: BAC ends

High quality sequence stp: 507.

Location/Qualifiers

1. 507

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="Plate=2266 Col=15 Row=A"

/sex="male"

/clone\_1lb="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelOAc11, BAC Clones in

E-Coli DH10B"

Query Match 19.1%; Score 33.8; DB 11; Length 507;

Best Local Similarity 53.4%; Pred. No. 7.6;

Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 5 TATGTGACGAGACGCCAGCTTCTGTGTTGTTACCGGACCGGACGACGACTCTTATC 64  
DB 106 TTTCCTTAGGACACCCATCATCTGTGTGGATGAGATCCGCTATACACCCCTTATG 165  
QY 65 GGACAGGAGCGCCCTCCATATCAGCGCGGCTTATCTCATGCGGTGACCGGACGAG 124  
DB 166 GGACAGGCGCTTCCACACGACTCTGCGCCAGAGAAACATGACGCGCATTTACACCAAT 225  
QY 125 GCGCGCGTCCCGC 137  
DB 226 GTCATCTCTCTC 238

## RESULT 3

LOCUS

BY342594

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 343)

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido,I., Osato,N., Saito,H., Yamana,K.,

Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,

Schondach,C., Gojodori,T., Balderelli,R., Hill,D.P., Bult,C.,

Hume,D.A., Quackenbush,J., Schriml,L.M., Kanpin,A., Matsuda,H.,

Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusan,V.,

Choitha,C., Corbani,L.E., Cousins,S., Dalla,E., Dragan,T.A.,

Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,

Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,

Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvie,E.D., Kanai,A.,

Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,

Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,

Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,

Numata,K., Okido,T., Pavan,W.J., Petosa,G., Pesole,G.,

Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S.,

Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,

Sandelin,A., Schneider,C., Sempke,C.A., Setou,M., Shimada,K.,

Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,

Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,

Wells,C., Wilming,L.G., Wrynshaw-Boris,A., Yanagisawa,M., Yang,I.,

Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,

Hayatsu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura,M.,

Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,Y., Aizawa,K.,

Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,

Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,

Shingawa,A., Yasunihi,A., Yoshino,M., Waterston,R., Lander,E.S.,

Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, [unr-http://genome.gsc.riken.jp/](http://unr-http://genome.gsc.riken.jp/)

Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,

Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,Y., Kono,H.,

Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,

Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vassil, Greece) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## ORIGIN

Query Match 19.0%; Score 33.6; DB 4; Length 343;  
Best Local Similarity 51.7%; Pred. No. 8.6;  
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 3 CTTATCTGACAGAGCCGCTTCTGTGTGTAACCGGACGCGAATCTCTTA 62  
Db 99 CTTAATGTGATCCCCCCCCCTTTCCTTCCCTAGCCTGACCCGAGAGCGCGTTT 158  
Qy 63 TCGGAACAGAGCGGCGCTCATATCAGCGCGGCTTATCTGATCGGCGTACCGGACG 122  
Db 159 CTCACACAGCTCGACACTCTCTCGGGGCGGACGCGCTCGCTCCCTGCGACACGG 218  
Qy 123 AGGCGCCGTCCTCCCTTATCGCGCC 147  
Db 219 GACAGAGCGTCCGCTTCCCTCTCC 243

RESULT 4  
LOCUS BY077721  
DEFINITION cDNA clone K630011C08 5', mRNA sequence.

ACCESSION BY077721  
VERSION BY077721.1 GI:26179135  
KEYWORDS Mus musculus  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS  
Ohtsaki, Y., Ohtsaki, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishikawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Hume, D.A., Quackenbush, J., Schrimm, L.W., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusa, V., Clouthier, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

TITLE  
JOURNAL PUBMED  
COMMENT  
Contact: Yoshinobu Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Aizawa, K., Akimura, T., Arikawa, T., Carrincci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
SOURCE  
1. 402  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K630011C08"  
/sex="male"  
/tissue\_type="liver"  
/dev\_stage="adult"  
/clone\_lib="RIKEN full-length enriched, adult male liver"

ORIGIN  
Query Match 19.0%; Score 33.6; DB 4; Length 402;  
Best Local Similarity 51.7%; Pred. No. 8.7;  
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
Qy 3 CTTATCTGACAGAGCCGCTTCTGTGTGTAACCGGACGCGAATCTCTTA 62

```

Db      89 CTTAATATGATCCCCCCCCCTTTCTTTCCCTAGCCGACGCGACAGCGCGTTT 148
Qy      63 TCGACACGAGACGCGCTCCATATCAGCCGCGCTTATCTCAGCGGTACCGGACAGC 122
Db      149 CTCACACGAGCTCGCACTCTCTCCGCGGCGGACGAGCTCGCGTCCCTCGCCACGANG 208
Qy      123 AGGCGCGCGCTCCGCTTATCGCGGC 147
Db      209 GACAGAGCTCCGCTTCCCTCC 233

RESULT 5
LOCUS   CV068875          931 bp      mRNA      linear      EST 24-AUG-2004
DEFINITION f2_new_chopped.fasta.Contig488 Preamplified custom cDNA library in
sequence.
ACCESSION CV068875
VERSION   CV068875.1 GI:51532039
KEYWORDS  EST.
SOURCE    Emiliania huxleyi
ORGANISM  Emiliania huxleyi
REFERENCE 1 Eukaryota; Haptophyceae; Isochrysidales; Emiliania.
AUTHORS   Wahland T.W., Zhang X. and Read B.A.
TITLE      Expressed Sequence Tag Profiles from Calcifying and Non-Calcifying
JOURNAL    Cultures of Emiliania huxleyi
COMMENT    Microaleontology (2004) In press
            Contact: Betsy Read
            Department of Biological Sciences
            California State University San Marcos
            333 S. Twin Oaks Valley Road, San Marcos, CA 92036-0001, USA
            Tel: 760 750 4129
            Email: bread@csusm.edu.

FEATURES
    source
        location/Qualifiers
            1..931
            /organism="Emiliania huxleyi"
            /strain="1516"
            /db_xref="taxon:2903"
            /dev_stage="late log growth phase"
            /clone_idb="Preamplified custom cDNA library in
            pcwvports6.1 (Reegen, Invitrogen Inc.)"
            /note="Emiliania huxleyi grown in Artificial Seawater
            (Guillard's F/2 media)."
```

```

REFERENCE 1 (bases 1 to 348)
AUTHORS   Sciurognathi; Muroidea; Muridae; Murinae; Mus.
            Okazaki, Y., Furuno, M., Kasukawa, T., Aachii, J., Bono, H., Kondo, S.,
            Nikaido, I., Osato, N., Saito, R., Suuki, H., Yamataka, I.,
            Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
            Schobach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Balt, C.,
            Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
            Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V.,
            Choctha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
            Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T.,
            Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
            Guerinich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
            Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
            Kurosch, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
            Matulis, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
            Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
            Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
            Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
            Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,
            Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
            Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
            Wells, C., Wilming, L.G., Wyszewski, A., Yanagisawa, M., Yang, I.,
            Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
            Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
            Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
            Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
            Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
            Shingawa, A., Yasunishi, A., Yoshino, M., Waterson, R., Lander, E.S.,
            Rogers, J., Birney, E. and Hayashizaki, Y.
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
            12466851
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
            Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
            Hirozane, T., Imotani, K., Ishii, Y., Itch, M., Kawai, J., Konno, H.,
            Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nunazaki, R.,
            Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
            Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
            Hayashizaki, Y. Direct Submision
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtration of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Tissues were provided by Vassilis Adinios (Biomedical Sciences
            Research Center "Al. Fleming" Institute of Immunology 14-16 Al.
            Fleming street 16672 Vathi, Greece) whose assistance we gratefully
            acknowledge.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
            Location/Qualifiers
                1..348
                /organism="Mus musculus"
                /mol_type="mRNA"
```

```

/db xref="taxon:10090"
/clone="L230014G21"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN
Query Match      18.6%; Score 33; DB 4; Length 348;
Best Local Similarity 51.7%; Pred. No. 13;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTATCGTGACAGACGCCAGCTTCTGTGTCTTAACCGACGCGACGCAACTCTTA 62
    |||||
DB 99 CTTATGTGATGCCCCCCCTTCTTCTTCGAGCTGACGGGACGAGCGCGCTT 158
    |||||

QY 63 TCGGACAGACAGCGCGCTTCATATCAGCGCGCTTATCTATGCGCTGACGACAGC 122
    |||||
DB 159 CCTACACAGAGCTCGACCTCTCTCGGGGGCGACGCGCTCGCTGCCCTGCACACAGC 218
    |||||

QY 123 AGGCGCCGCTCCGCTTATCGCGC 147
    |||||
DB 219 GACAGAGCGTCCGCTTCTCTTC 243
    |||||

RESULT 7
BY337247 353 bp mRNA linear EST 11-DEC-2002
LOCUS BY337247 RIKEN full-length enriched, whole joints Mus musculus cDNA
DEFINITION clone L230006D15 5', mRNA sequence.
ACCESSION BY337247
VERSION BY337247.1 GI:26532790
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 353)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, R., Otsu, N., Saito, R., Suzuki, H., Yamana, T.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schiraldi, L. M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V.,
Chothia, C., Corbani, L. E., Cousins, S., Dalli, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
Gardbodi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, T.,
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawai, H., Kawasawa, Y., Kedzierzki, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Matsui, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W. J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. O., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wyntaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E., and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
JOURNAL PUBMED
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

```

```

Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,
Miyazaki, A., Murata, M., Nakamura, K., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vassil, Greece) whose assistance we gratefully
acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES
SOURCE
1..353
location/qualifiers
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L230006D15"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN
Query Match      18.6%; Score 33; DB 4; Length 353;
Best Local Similarity 51.7%; Pred. No. 13;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 3 CTATCGTGACAGACGCCAGCTTCTGTGTCTTAACCGACGCGACGCAACTCTTA 62
    |||||
DB 99 CTTATGTGATGCCCCCCCTTCTTCTTCGAGCTGACGGGACGAGCGCGCTT 158
    |||||

QY 63 TCGGACAGACAGCGCGCTTCATATCAGCGCGCTTATCTATGCGCTGACGACAGC 122
    |||||
DB 159 CCTACACAGAGCTCGACCTCTCTCGGGGGCGACGCGCTCGCTGCCCTGCACACAGC 218
    |||||

QY 123 AGGCGCCGCTCCGCTTATCGCGC 147
    |||||
DB 219 GACAGAGCGTCCGCTTCTCTTC 243
    |||||

RESULT 8
BY013584 354 bp mRNA linear EST 06-DEC-2002
LOCUS BY013584 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus
DEFINITION musculus cDNA clone G730040E06 5', mRNA sequence.
ACCESSION BY013584
VERSION BY013584.1 GI:26073833
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 354)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

```

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, K. I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schirml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Bruscia, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gerbaldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, R. M., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, I. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Naitas, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Varrault, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyszynski, B. S., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Akakawa, H., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

JOURNAL  
PUBMED  
COMMENT

## TITLE

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES  
source

Location/Qualifiers  
1. 354  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="G730040E06"  
/tissue\_type="lung"  
/cell\_line="RCB-0558 LLC"  
/clone\_lhb="RIKEN full-length enriched, lung RCB-0558 LLC cDNA"

ORIGIN

Query Match 18.6%; Score 33; DB 4; Length 354;  
Best Local Similarity 51.7%; Pred. No. 13;  
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 3 CTTATGCTGACAGACGCCAGCTTCTGTGTGTCCTAACCGCAGCGACCACTCTTA 62  
Db 116 CTTAATGTGATCCCCCCCCCTTCTCTTCGTAAGCTGACGCGGACGACCGCTT 175

Qy 63 TCGGACAGACCGCCTCATATCAGCGCGGTATCTCAGCCCGTACCGGACG 122

Db 176 CCGACAGCTGACCTCTCTCGGAGCGCAGCGCTCGCCCTGCCCCACACG 235

Qy 123 AGGCGCCGTCCTTATCGGCC 147

Db 236 GACAGACGCTCCGCTTCCTTC 260

RESULT 9  
BY340467  
LOCUS

DEFINITION  
BY340467 RIKEN full-length enriched, whole joints Mus musculus cDNA

ACCESSION  
BY340467

VERSION  
BY340467.1 GI:26569955

KEYWORDS  
EST.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus

REFERENCE  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 355)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, T., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, K. I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schombach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C.,

Hume, D. A., Quackenbush, J., Schirml, L. M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Bruscia, V.,

Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,

Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,

Gerbaldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,

Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Nomura, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,

Varrault, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L. G., Wyszynski, B. S., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, H., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H.,





Matches	75; Conservative	0; Mismatches	70; Indels	0; Gaps	0;
Qy	3 CTTATCTGTGACAGAGCGCCAGCTTCTCTGTGTGTCTTAACCGAGCGGACGAACTCTTA 62				
Db	99 CTTAATGTATGATCCCCCCCCCTTTCTTTCTTCCGTACCTGACCGGACAGACGCGCTTT 158				
Qy	63 TCGGAACAGACAGCGCGCTTCATATCAGCCGCGCTTATCTCATCGGCGTGAACGACAG 122				
Db	159 CTCACACAGCTCGACACTCTCTCGGGGGCGACCGCTCGCCTCCCTGCTCCACACAGG 218				
Qy	123 AGCGCGCCGCTCCGCTTATCGCGCC 147				
Db	219 GACAGAGCGTCCGCTTCCTCTCC 243				
RESULT 11	BY339252 362 bp mRNA linear EST 12-DEC-2002				
BY339252	BY339252 RIKEN full-length enriched, whole joints Mus musculus cDNA				
DEFINITION	clone l230021B07 5', mRNA sequence.				
ACCESSION	BY339252				
VERSION	BY339252.1 GI:26568740				
KEYWORDS	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 362) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Ryohbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brunic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustinch, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszynski, B., Yang, Y., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Atake, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, B. and Hayashizaki, Y.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
PUBMED	12466851				
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and				

FEATURES	source				
LOCATION/Qualifiers	1..362				
ORGANISM	/organism="Mus musculus"				
DEFINITION	/db_xref="taxon:10090"				
ACCESSION	/clone="l230021B07"				
VERSION	/tissue_type="whole joints"				
KEYWORDS	/clone_1ib="RIKEN full-length enriched, whole joints"				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Hayashizaki, Y. Direct Submision				
AUTHORS	Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)				
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)				
	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexed sequencer. Genome Res. 10 (11), 1757-1771 (2000)				
	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)				
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.				
	Tissues were provided by Vassilis Aidinis ( Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece ) whose assistance we gratefully acknowledge.				
	Please visit our web site (http://genome.gsc.riken.go.jp) for further details.				
Query Match	18.6%	Score 33;	DB 4;	Length 362;	
Best Local Similarity	51.7%	Pred. No. 13;			
Matches	75; Conservative	0; Mismatches	70; Indels	0; Gaps	0;
Qy	3 CTTATCTGTGACAGAGCGCCAGCTTCTCTGTGTGTCTTAACCGAGCGGACGAACTCTTA 62				
Db	99 CTTAATGTATGATCCCCCCCCCTTTCTTTCTTCCGTACCTGACCGGACAGACGCGCTTT 158				
Qy	63 TCGGAACAGACAGCGCGCTTCATATCAGCCGCGCTTATCTCATCGGCGTGAACGACAG 122				
Db	159 CTCACACAGCTCGACACTCTCTCGGGGGCGACCGCTCGCCTCCCTGCTCCACACAGG 218				
Qy	123 AGCGCGCCGCTCCGCTTATCGCGCC 147				
Db	219 GACAGAGCGTCCGCTTCCTCTCC 243				
RESULT 12	BY083658 371 bp mRNA linear EST 07-DEC-2002				
BY083658	BY083658 RIKEN full-length enriched, 10 days neonate heart Mus musculus cDNA clone K630044L13 5', mRNA sequence.				
DEFINITION	musculus cDNA clone K630044L13 5', mRNA sequence.				
ACCESSION	BY083658				
VERSION	BY083658.1 GI:26203394				
KEYWORDS	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 371) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Ryohbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brunic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S.,				



Guellinich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kani, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Matsuda, L., Marchionni, L., McKenzie, L., Miki, H., Nagishima, T., Numa, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Velasco, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Maki, K., Kawaji, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

# JOURNAL PUBMED

1246681

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saitoh-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

## FEATURES source

Location/Qualifiers  
1. 371  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RG300441.3"  
/tissue\_type="heart"  
/dev\_stage="10 days neonate"  
/clone\_lib="RIKEN full-length enriched, 10 days neonate heart"

## ORIGIN

Query Match 18.6%; Score 33; DB 4; Length 371;  
Best Local Similarity 51.7%; Pred. No. 13;  
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
3 CTTATCGTACGAGCGCGCTTCCTGTTCTACCGCGAGCGCACTCTTA 62  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 111 CTTATGATGATCCCGCCCTTTCTTCTCGGTACGCGGCGGACGCGCGCTT 170  
Cy 63 TCGAAGCAGACGGCGCTCCATATGACCGCGGTATTCATGCGGTACCGACG 122  
Db 171 CTTACACGAGCTCGCATCTCTCGGGGCGGCGGCGCTCGCGCTCCACACG 230  
Cy 123 AGGCGCGCGCGCGCGCTTATGCGGCG 147  
Db 231 GACAGAGCTTCGCGCTTCCTCTCC 255

## RESULT 13

## CP916169

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL PUBMED

## COMMENT

Genome Res. 11 (9), 1553-1558 (2001)  
11544199  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@nigun.gsc.nih.gov  
Plate: B0991 row: A column: 12  
Seq primer: M13 Reverse  
High quality sequence stop: 611  
POLYA=No.

## FEATURES source

Location/Qualifiers  
1. 611  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="nlaEST:B0991A12-5"  
/db\_xref="taxon:10090"  
/clone="NIA:B0991A12 IMAGE:30480971"  
/dev\_stage="Unfertilized Egg"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse Unfertilized Egg cDNA Library (Long 1)"  
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.gsc.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were extracted from a pool of 1488 unfertilized eggs. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]. 5'-pGACTAGTTCATGATCGAGCGCGCGCTTTT-3', treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The



average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

Query Match	18.6%	Score 33	DB 5	Length 624
Best Local Similarity	51.7%	Pred. No. 14		
Matches 75; Conservative	0	Mismatches 70	Indels 0	Gaps 0

Oy 3 CTTATCGGACAGAGAGCGCAGCTTCCTGTGTTGCTAACCGCAGCCGAGCACTCTTAA 62  
 Db 33 CTTAATGTGATCCCCCCCCCTTCTCTTCTCGTAGCTGACGCGGACACCCGGTTT 92  
 Oy 63 TCGGAACAGAGAGCGCCTCATATCAGCGCGCGGTATCTCATGCGCGTACCGGACACG 122  
 Db 93 CCTACACGACTGCACTCTCTCGGGGGGCGACAGGTCGCGCTTGCCCCCTGCCCCACACG 152  
 Oy 123 AGCGCCCGTCCCGCTTATCGCGCC 147  
 Db 153 GACAGAGCGTCCCGCTTCCCTTCC 177

Search completed: June 15, 2006, 18:15:44  
Job time : 1485.21 secs

**THIS PAGE BLANK (USPTO)**



```

; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2371
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-2371

Query Match
Best Local Similarity 55.2%; Score 29.8; DB 3; Length 1311;
Matches 58; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 GCGGACGCAACTCTTATGGAACAGACGCGCTCCATATACGCGCGGTATCTCA 104
    |||||
DB 368 GCGGACGCTCTCTGCTGTGACCGGAGACATCTCATCTTCTGACGCGTCAACC 427
    |||||

QY 105 TGCGCGTGACCGGACAGAGCGCGCTCCGCTTATCGCGCTTA 149
    |||||
DB 428 GGAACGGGACATTGACGCGCGCGGTCACTTTATCCCA 472
    |||||

RESULT 3
US-09-902-540-1101/c
; Sequence 1101, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1101
; LENGTH: 13579
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1101

Query Match
Best Local Similarity 55.2%; Score 29.8; DB 3; Length 13579;
Matches 58; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 GCGGACGCAACTCTTATGGAACAGACGCGCTCCATATACGCGCGGTATCTCA 104
    |||||
DB 1063 GCGGACGCTCTCTGCTGTGACCGGAGACATCTCATCTTCTGACGCGGTCAACC 1004
    |||||

QY 105 TGCGCGTGACCGGACAGAGCGCGCTCCGCTTATCGCGCTTA 149
    |||||
DB 1003 GGAACGGGACATTGACGCGCGCGGTCACTTTATCCCA 959
    |||||

RESULT 4
US-09-252-991A-12370/c
; Sequence 12370, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
```

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12370
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12370

Query Match
Best Local Similarity 53.0%; Score 28.6; DB 3; Length 585;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 8 CGTACAGACGCGCCTTCTGTGTGCTTAACGCGACGCGGACGCAACTCTTATCGGA 67
    |||||
DB 534 CGTGCCTGTGCGACAGGCGCTGTGTGCGACGCGGATGCCCTTGATACCGTGTGCA 475
    |||||

QY 68 ACAGACGCGCCTTCATATACGCGCGCTTATCTATGCGCTGACGCGACAG 122
    |||||
DB 474 GCTGACGCGCGCGGACATCAAGCGGACGCTGAGGCGACGCTCCCGGAATG 420
    |||||

RESULT 5
US-09-252-991A-12097
; Sequence 12097, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12097
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12097

Query Match
Best Local Similarity 53.0%; Score 28.6; DB 3; Length 1443;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 8 CGTACAGACGCGCCTTCTGTGTGCTTAACGCGACGCGGACGCAACTCTTATCGGA 67
    |||||
DB 810 CGTGCCTGTGCGACAGGCGCTGTGTGCGACGCGGATGCCCTTGATACCGTGTGCA 869
    |||||

QY 68 ACAGACGCGCCTTCATATACGCGCGCTTATCTATGCGCTGACGCGACAG 122
    |||||
DB 870 GCTGACGCGCGGACATCAAGCGGACGCTGAGGCGACGCTCCCGGAATG 924
    |||||

RESULT 6
US-09-252-991A-12323/c
; Sequence 12323, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12323
```

LENGTH: 1551  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12323

Query Match 16.2%; Score 28.6; DB 3; Length 1551;  
Best Local Similarity 53.0%; Pred. No. 12;  
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
QY 8 CGTACAGAGCGCCCTCTCTGTTCTTAACCGGACCGGACGCACTCTTATCGGA 67  
DB 715 CGTGCCTGCTGGACGGGGCTGTGTCGACCGCGATGCCGCTTCGATACCGTGTGCGA 656  
QY 68 ACAGAGCGCGCCCTCATATCAGCGCGGCTTATCTCATGCGGACCGGACGAG 122  
DB 655 GCTGACCGCGCCGACGATCAAGCCGACGCTGAGGCGACGTGCGGAAATG 601

RESULT 7  
US-09-252-991A-12237  
Sequence 12237, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 12237  
LENGTH: 2085  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-12237

Query Match 16.2%; Score 28.6; DB 3; Length 2085;  
Best Local Similarity 53.0%; Pred. No. 12;  
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
QY 8 CGTACAGAGCGCCCTCTCTGTTCTTAACCGGACCGGACGCACTCTTATCGGA 67  
DB 802 CGTGCCTGCTGGACGGGGCTGTGTCGACCGCGATGCCGCTTCGATACCGTGTGCGA 861  
QY 68 ACAGAGCGCGCCCTCATATCAGCGCGGCTTATCTCATGCGGACCGGACGAG 122  
DB 862 GCTGACCGCGCCGACGATCAAGCCGACGCTGAGGCGACGTGCGGAAATG 916

RESULT 8  
US-08-899-437-22  
Sequence 22, Application US/08899437  
Patent No. 6121415  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,437  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Deirdre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2091 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: Human NRG3B2 (hNRG3B2)  
LOCATION: 1-2091  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
US-08-899-437-22

Query Match 16.2%; Score 28.6; DB 3; Length 2091;  
Best Local Similarity 58.0%; Pred. No. 12;  
Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;  
QY 55 ACTCCTTATCGGAACAGAGCGCGCTTCATATCAGCGCGGCTTATCTCATGCGCGGAG 114  
DB 448 ACCGCCAACCGGATTAGACTCTCCCTGACCAACA-TCAGCGGCGCGCCACTGCTTCCC 506  
QY 115 CGAGACGAGCGCGCCCTCTCCGCTTATTCGCGCTTAAATACAGCCCGCAAGCATGTG 173  
DB 507 CGGACACCGGGTCCCATCCGCGGACCGCGCTCCACCAAGACGGAACACTGCGG 565

RESULT 9  
US-09-126-121-22  
Sequence 22, Application US/09126121  
Patent No. 6252051  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
TITLE OF INVENTION: Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,121  
FILING DATE: 30-Jul-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Deirdre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2091 base pairs

TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
FEATURE: Human NR3B2 (hNR3B2)  
LOCATION: 1-2091  
IDENTIFICATION METHOD:  
OTHER INFORMATION:

US-09-126-121-22

Query Match 16.2%; Score 28.6; DB 3; Length 2091;  
Best Local Similarity 58.0%; Pred. No. 12;

Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 55 ACTCCTTATCGGAACAGAGCGCGCTGCTTATCGAGCGCGCTTATCGAGCGCGCGTGG 114  
DB 448 AGCCCAACCGGATTAACACTCGCTTACCAACCA-TCACGCGGCGCGCGCTTCC 506  
QY 115 CGGACAGAGCGCGCGCTTATCGAGCGCGCTTATCGAGCGCGCGTGG 173  
DB 507 CGGGACCGGGTGCCTATCCGGCGAGCGCGCGCTTACCAACAGAGAGCACTGGCG 565

RESULT 10

US-09-107-979-22

Sequence 22, Application US/09107979  
Patent No. 6994856

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPac (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107, 979

FILING DATE: 30-Jun-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Deirdre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 2091 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NR3B2 (hNR3B2)

LOCATION: 1-2091

IDENTIFICATION METHOD:

OTHER INFORMATION:

Query Match 16.2%; Score 28.6; DB 4; Length 2091;  
Best Local Similarity 58.0%; Pred. No. 12;

Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 55 ACTCCTTATCGGAACAGAGCGCGCTTATCGAGCGCGCTTATCGAGCGCGTGG 114

DB 448 AGCCCAACCGGATTAACACTCGCTTACCAACCA-TCACGCGGCGCGCGCTTCC 506  
QY 115 CGGACAGAGCGCGCGCTTATCGAGCGCGCTTATCGAGCGCGCGTGG 173  
DB 507 CGGGACCGGGTGCCTATCCGGCGAGCGCGCGCTTACCAACAGAGAGCACTGGCG 565

RESULT 11

US-08-899-437-5

Sequence 5, Application US/08899437  
Patent No. 6121415

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPac (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,437

FILING DATE: 24-Jul-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Conley, Deirdre L.

REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2502 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

FEATURE:

NAME/KEY: Human NR3B1 (hNR3B1)/nucleic acid seq.

LOCATION: 1-2502

IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-899-437-5

Query Match 16.2%; Score 28.6; DB 3; Length 2502;  
Best Local Similarity 58.0%; Pred. No. 13;

Matches 69; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 55 ACTCCTTATCGGAACAGAGCGCGCTTATCGAGCGCGCTTATCGAGCGCGTGG 114  
DB 591 AGCCCAACCGGATTAACACTCGCTTACCAACCA-TCACGCGGCGCGCGCTTCC 649  
QY 115 CGGACAGAGCGCGCGCTTATCGAGCGCGCTTATCGAGCGCGCGTGG 173  
DB 650 CGGGACCGGGTGCCTATCCGGCGAGCGCGCGCTTACCAACAGAGAGCACTGGCG 708

RESULT 12

US-09-126-121-5

Sequence 5, Application US/09126121  
Patent No. 6252051

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao

TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related





## RESULT 15

```
US-09-949-016-13983/c
; Sequence 13983, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ. ID NO 13983
; LENGTH: 57320
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(57320)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13983
```

Query Match 16.0%; Score 28.4; DB 3; Length 57320;

Best Local Similarity 49.3%; Pred. No. 31; Matches 74; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```
QY 15 GGACGCCAGCTTCTGTGTTGTTACTACCGCAGCCGAGCAACTCTTATCGAACAGAGAC 74
DB 2034 GGAGACCGCATCTTCTGTCAGACCCGAGCGCCGCGACGACGACGACGACGACCGTC 1975
QY 75 GCGCCTTCATATACAGCCGCGCTTATCTCATGCGCGGTGACCGAGACGAGCGCGCTCC 134
DB 1974 GCGCGTCACCTTCGCGGAGCGCGCGCGCTGCAAGCGTCACAGCGGACTCCGCCCGCGG 1915
QY 135 CGCTTATCGCGCTTATTAATACAGCCCGCA 164
DB 1914 GCGCGGGGCGACCCGAGGAGACAGCGGCA 1885
```

Search completed: June 15, 2006, 17:18:11  
Job time : 62.6721 secs



```
RESULT 2
US-10-622-088-126
; Sequence 126, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942.545007
; CURRENT APPLICATION NUMBER: US/10/622,088
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: OpIe2 promoter sequence
US-10-622-088-126

Query Match      100.0%; Score 177; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 3e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTATCGTAGACGAGCGCAGCTTCTGTGTGCTAACGCGACGCGAAGTCTCT 60
DB      355 GTCTTATCGTAGACGAGCGCAGCTTCTGTGTGCTAACGCGACGCGAAGTCTCT 414
QY      61  TATCGAACAAGACGCGCCTTCATATCAGCGCGCGTTATCTCATYGGCGTGAACGAGCA 120
DB      415 TATCGAACAAGACGCGCCTTCATATCAGCGCGCGTTATCTCATYGGCGTGAACGAGCA 474
QY      121 CGAGGCGCGCGTCCCGCTTATCGGCGCTTAATAATACGCGCGAAGATCTGGTAAA 177
DB      475 CGAGGCGCGCGTCCCGCTTATCGGCGCTTAATAATACGCGCGAAGATCTGGTAAA 531

RESULT 3
US-09-896-888a-1
; Sequence 1, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Insect Expression Vectors
; FILE REFERENCE: 80021-44
; CURRENT APPLICATION NUMBER: US/09/896,888A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/048,911
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/049,946
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
```

```
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Orygia pseudotsugata
US-09-896-888a-1

Query Match      100.0%; Score 177; DB 3; Length 564;
Best Local Similarity 100.0%; Pred. No. 3e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTATCGTAGACGAGCGCAGCTTCTGTGTGCTAACCGGACCGGACGCAATCTCT 60
DB      351 GTCTTATCGTAGACGAGCGCAGCTTCTGTGTGCTAACCGGACCGGACGCAATCTCT 410
QY      61  TATCGAACAAGACGCGCCTTCATATCAGCGCGGCTTAATCATGCGGCTGACCGGACA 120
DB      411 TATCGAACAAGACGCGCCTTCATATCAGCGCGGCTTAATCATGCGGCTGACCGGACA 470
QY      121 CGAGGCGCGCGTCCCGCTTATCGGCGCTTAATAATACGCGCGAAGATCTGGTAAA 177
DB      471 CGAGGCGCGCGTCCCGCTTATCGGCGCTTAATAATACGCGCGAAGATCTGGTAAA 527

RESULT 4
US-10-295-074-60
; Sequence 60, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2Zop2f expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: AvalI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaLI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
```

NAME/KEY: misc\_recomb  
LOCATION: (2204)..(2209)  
OTHER INFORMATION: NcoI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2284)..(2289)  
OTHER INFORMATION: AvalI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2294)..(2299)  
OTHER INFORMATION: AvalI, SmaI, and XmaI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2551)..(2556)  
OTHER INFORMATION: ApaLI site  
US-10-295-074-60

Query Match 100.0%; Score 177; DB 7; Length 2773;  
Best Local Similarity 100.0%; Pred. No. 3.8e-54;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGAAGAGCGCGAGCTTCTGTGTGCTTAACCGAGCCGAGCAACTCTCT 60  
DB 355 GTCTTATCGTGAAGAGCGCGAGCTTCTGTGTGCTTAACCGAGCCGAGCAACTCTCT 414  
QY 61 TATCGGAACAGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGCTGACCGACA 120  
DB 415 TATCGGAACAGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGCTGACCGACA 474  
QY 121 CGAGGGCCCGTCCCGCTTATCGCGCTTAAATATACGCCGCAAGATCTGTATAA 177  
DB 475 CGAGGGCCCGTCCCGCTTATCGCGCTTAAATATACGCCGCAAGATCTGTATAA 531

RESULT 5  
US-10-846-911-60  
Sequence 60, Application US/10846911  
Publication No. US20040258660A1  
GENERAL INFORMATION:  
APPLICANT: KLYSNER, Steen  
APPLICANT: NIELSEN, Finn Stausholm  
APPLICANT: BRATT, Tomae  
APPLICANT: VOLDORF, Bjorn  
TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS  
FILE REFERENCE: 674542-2018  
CURRENT FILING DATE: 2004-05-14  
PRIOR APPLICATION NUMBER: US/10/846, 911  
PRIOR FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: PCT/DK02/00764  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: PA 2001 01702  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 60  
LENGTH: 2773  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: p2ZOP2F expression vector for insect cells  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (561)..(566)  
OTHER INFORMATION: HindIII site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (573)..(578)  
OTHER INFORMATION: AvalI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (586)..(591)

OTHER INFORMATION: EcoRI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (593)..(598)  
OTHER INFORMATION: BamHI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (625)..(630)  
OTHER INFORMATION: ClaI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (629)..(634)  
OTHER INFORMATION: ClaI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (1156)..(1161)  
OTHER INFORMATION: ApaLI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2128)..(2133)  
OTHER INFORMATION: PstI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2204)..(2209)  
OTHER INFORMATION: NcoI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2284)..(2289)  
OTHER INFORMATION: NcoI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2294)..(2299)  
OTHER INFORMATION: AvalI, SmaI, and XmaI site  
FEATURE:  
NAME/KEY: misc\_recomb  
LOCATION: (2551)..(2556)  
OTHER INFORMATION: ApaLI site  
US-10-846-911-60

Query Match 100.0%; Score 177; DB 9; Length 2773;  
Best Local Similarity 100.0%; Pred. No. 3.8e-54;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTATCGTGAAGAGCGCGAGCTTCTGTGTGCTTAACCGAGCCGAGCAACTCTCT 60  
DB 355 GTCTTATCGTGAAGAGCGCGAGCTTCTGTGTGCTTAACCGAGCCGAGCAACTCTCT 414  
QY 61 TATCGGAACAGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGCTGACCGACA 120  
DB 415 TATCGGAACAGAGCGCGCTTCATATCAGCGCGCTTATCTCATGCGCTGACCGACA 474  
QY 121 CGAGGGCCCGTCCCGCTTATCGCGCTTAAATATACGCCGCAAGATCTGTATAA 177  
DB 475 CGAGGGCCCGTCCCGCTTATCGCGCTTAAATATACGCCGCAAGATCTGTATAA 531

RESULT 6  
US-10-939-107-60  
Sequence 60, Application US/10939107  
Publication No. US20050180947A1  
GENERAL INFORMATION:  
APPLICANT: Pedersen, Hans Rudolf  
APPLICANT: Ederl, Bjarke  
APPLICANT: Pedersen, Louise Henriette  
APPLICANT: Rasmussen, Peter Birk  
TITLE OF INVENTION: Novel Application of Vaccination Against TNF-alpha  
FILE REFERENCE: 674542-2020  
CURRENT FILING DATE: 2004-09-10  
PRIOR APPLICATION NUMBER: PCT/DK03/00147  
PRIOR FILING DATE: 2003-03-11  
PRIOR APPLICATION NUMBER: 60/363,128  
PRIOR FILING DATE: 2002-03-11

```

; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 2773
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p2ZOp2F expression vector for insect cells
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (561)..(566)
; OTHER INFORMATION: HindIII site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (573)..(578)
; OTHER INFORMATION: AatI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (586)..(591)
; OTHER INFORMATION: EcoRI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (593)..(598)
; OTHER INFORMATION: BamHI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (625)..(630)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (629)..(634)
; OTHER INFORMATION: ClaI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (1156)..(1161)
; OTHER INFORMATION: ApaLI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2128)..(2133)
; OTHER INFORMATION: PstI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2204)..(2209)
; OTHER INFORMATION: NcoI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2284)..(2289)
; OTHER INFORMATION: AatI site
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: (2294)..(2299)
; OTHER INFORMATION: AatI, SmaI, and XbaI site
; NAME/KEY: misc_recomb
; LOCATION: (2551)..(2556)
; OTHER INFORMATION: ApaLI site
US-10-939-107-60

Query Match      100.0%; Score 177; DB 10; Length 2773;
Best Local Similarity 100.0%; Pred. No. 3,8e-54;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 7
US-10-622-088-89
; Sequence 89, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942,5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,335
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/398,617
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/427,231
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 60/456,496
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 60/474,940
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 5038
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PIB/V5-His-DEST
US-10-622-088-89

Query Match      99.1%; Score 175.4; DB 9; Length 5038;
Best Local Similarity 99.4%; Pred. No. 1.6e-53;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

RESULT 8
US-10-622-088-127
; Sequence 127, Application US/10622088
; Publication No. US20040219516A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, Robert P.
; APPLICANT: Welch, Peter J.
; APPLICANT: Harwood, Steven
; APPLICANT: Madden, Knut
; APPLICANT: Frimpong, Kenneth
; APPLICANT: Franke, Kenneth E.
; TITLE OF INVENTION: Viral Vectors Containing Recombination Sites
; FILE REFERENCE: 0942,5450007
; CURRENT APPLICATION NUMBER: US/10/622,088
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: PCT/US03/22437
; PRIOR FILING DATE: 2003-07-18

```

PRIOR APPLICATION NUMBER: US 60/396,335  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: US 60/398,617  
PRIOR FILING DATE: 2002-07-26  
PRIOR APPLICATION NUMBER: US 60/427,231  
PRIOR FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: US 60/456,496  
PRIOR FILING DATE: 2003-03-24  
PRIOR APPLICATION NUMBER: US 60/474,940  
PRIOR FILING DATE: 2003-06-03  
NUMBER OF SEQ ID NOS: 165  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 127  
LENGTH: 147  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Recombination region of pIB/V5-His-DEST  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (141)..(148)  
OTHER INFORMATION: n may be any nucleotide  
US-10-622-088-127

Query Match 23.2%; Score 41; DB 9; Length 147;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGTAA 177  
Db 1 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGTAA 41

RESULT 9  
US-10-622-088-149  
Sequence 149, Application US/10622088  
Publication No. US20040219516A1  
GENERAL INFORMATION:  
APPLICANT: Bennett, Robert P.  
APPLICANT: Welch, Peter J.  
APPLICANT: Harwood, Steven  
APPLICANT: Madden, Knut  
APPLICANT: Frimpong, Kenneth  
APPLICANT: Franke, Kenneth B.  
TITLE OR INVENTION: Viral Vectors Containing Recombination Sites  
FILE REFERENCE: 0942.545007  
CURRENT APPLICATION NUMBER: US/10/622,088  
CURRENT FILING DATE: 2003-07-18  
PRIOR APPLICATION NUMBER: PCT/US03/22437  
PRIOR FILING DATE: 2003-07-18  
PRIOR APPLICATION NUMBER: US 60/396,335  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: US 60/398,617  
PRIOR FILING DATE: 2002-07-26  
PRIOR APPLICATION NUMBER: US 60/427,231  
PRIOR FILING DATE: 2002-11-19  
PRIOR APPLICATION NUMBER: US 60/456,496  
PRIOR FILING DATE: 2003-03-24  
PRIOR APPLICATION NUMBER: US 60/474,940  
PRIOR FILING DATE: 2003-06-03  
NUMBER OF SEQ ID NOS: 165  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 149  
LENGTH: 325  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Recombination region of pIB/V5 His DEST  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (141)..(142)  
OTHER INFORMATION: n is a, c, g, or t  
FEATURE:

NAME/KEY: CDS  
LOCATION: (145)..(276)  
US-10-622-088-149

Query Match 23.2%; Score 41; DB 9; Length 325;  
Best Local Similarity 100.0%; Pred. No. 0.00018;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGTAA 177  
Db 1 CTTATCGCGCCTTAATAATACAGCCGCAACGATCTGTAA 41

RESULT 10  
US-10-156-761-2538/C  
Sequence 2538, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OR INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 2538  
LENGTH: 3138  
TYPE: DNA  
ORGANISM: Streptomyces avermitilis  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(3138)  
US-10-156-761-2538

Query Match 18.2%; Score 32.2; DB 7; Length 3138;  
Best Local Similarity 61.2%; Pred. No. 0.44;  
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 CTTATCGAAGAGAGAGCGCTCCATATCAGCGCGGCTTATCTCATGCGGCGTGAACCG 117  
Db 2575 CTTATCGAAGAGAGAGCGCTCCATATCAGCGCGGCTTATCTCATGCGGCGTGAACCG 2516  
QY 118 ACAAGAGAGCGCGCTCCATATC 142  
Db 2515 CTTAGAGGCGCTCCATATC 2491

RESULT 11  
US-10-156-761-1  
Sequence 1, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OR INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109  
SEQ ID NO 1  
LENGTH: 9025608  
TYPE: DNA  
ORGANISM: Streptomyces avermitilis  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (418715)  
OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Query Match 18.2%; Score 32.2; DB 7; Length 9025608;  
Best Local Similarity 61.2%; Pred. No. 1.6;  
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 CCTATCGAAGACGCGCTTCATATCAGCGCGTTATCTCATGGCGGTGACCGG 117  
DB 3129487 CCTCCACCGAACCGAGCGCTGCCCTTGATGATGAGTTGCTTCCGACCGACCGG 3129546  
QY 118 ACAAGAGCGCGCCGCTCCCGCTTATC 142  
DB 3129547 CCTTGAGGCGCTGCTCCAGGTTCTC 3129571

RESULT 12  
US-10-779-543-8561/c  
Sequence 8561, Application US/10779543  
Publication No. US2005022917A1  
GENERAL INFORMATION:  
APPLICANT: Williams et al  
TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED  
FILE REFERENCE: 2300-21302  
CURRENT APPLICATION NUMBER: US/10/779,543  
CURRENT FILING DATE: 2004-02-12  
PRIOR APPLICATION NUMBER: 10/076,555  
PRIOR FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: 09/217,471  
PRIOR FILING DATE: 1998-12-21  
PRIOR APPLICATION NUMBER: 60/068,755  
PRIOR FILING DATE: 1997-12-23  
PRIOR APPLICATION NUMBER: 60/080,664  
PRIOR FILING DATE: 1998-04-03  
PRIOR APPLICATION NUMBER: 60/105,234  
PRIOR FILING DATE: 1998-10-21  
PRIOR APPLICATION NUMBER: 09/297,648  
PRIOR FILING DATE: 2000-04-10  
PRIOR APPLICATION NUMBER: PCT/US99/01619  
PRIOR FILING DATE: 1999-01-28  
PRIOR APPLICATION NUMBER: 60/072,910  
PRIOR FILING DATE: 1998-01-28  
PRIOR APPLICATION NUMBER: 60/075,954  
PRIOR FILING DATE: 1998-02-24  
PRIOR APPLICATION NUMBER: 60/080,114  
PRIOR FILING DATE: 1998-03-31  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 23767  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8561  
LENGTH: 921  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 4, 6, 29, 88, 91, 146, 218, 378, 413, 482, 485, 500, 508,  
LOCATION: 531, 573, 585, 623, 637, 642, 651, 662, 697, 704, 724,  
LOCATION: 731, 778, 835, 837, 839, 842, 852, 868, 882, 884, 885, 891,  
LOCATION: 892, 899, 901  
OTHER INFORMATION: n = A,T,C or G  
US-10-779-543-8561

Query Match 18.1%; Score 32; DB 10; Length 921;  
Best Local Similarity 51.5%; Pred. No. 0.43;  
Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 16 GAGCCAGCTTCTGTGTGTTAACCAGCGAGCGAAGTCTTATCGAAGACGAG 75  
DB 157 GCGAGCTGGGCGNCCAGAGTGAGTGTCTCCAGCGCGGTCTCCGCGCTGCCGGTGGG 98  
QY 76 CGCTCCATATCAGCGCGCGTTATCTCATGCGCGGTGACCGGACGAGCGCGCTCC 135  
DB 97 GGGGTGCAANCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 38  
QY 136 GCTTATCGGCGCC 147  
DB 37 TAAATCGAGGCC 26

RESULT 13  
US-11-097-143-25442  
Sequence 25442, Application US/11097143  
Publication No. US20050208558A1  
GENERAL INFORMATION:  
APPLICANT: Venter, J. Craig  
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
FILE REFERENCE: CL000728  
CURRENT APPLICATION NUMBER: US/11/097,143  
CURRENT FILING DATE: 2005-04-04  
PRIOR APPLICATION NUMBER: 60/157,832  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: 60/160,191  
PRIOR FILING DATE: 1999-10-19  
PRIOR APPLICATION NUMBER: 60/161,932  
PRIOR FILING DATE: 1999-10-28  
PRIOR APPLICATION NUMBER: 60/164,769  
PRIOR FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: 60/173,383  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: 60/175,693  
PRIOR FILING DATE: 2000-01-12  
PRIOR APPLICATION NUMBER: 60/184,831  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: 60/191,637  
PRIOR FILING DATE: 2000-03-23  
NUMBER OF SEQ ID NOS: 43008  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 25442  
LENGTH: 1268  
TYPE: DNA  
ORGANISM: DROSOPHILA  
US-11-097-143-25442  
Query Match 17.4%; Score 30.8; DB 13; Length 1268;  
Best Local Similarity 55.7%; Pred. No. 1.2; 47; Indels 0; Gaps 0;  
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 17 AGCGAGCTTCTGTGTGTTAACCAGCGAGCGAAGTCTTATCGAAGACGAGCG 76  
DB 657 ACAAGAGCTTCTCGGTGAGAGTGCCGACAGCGGGTCTTGTCTTGGCGGCTCCGGTGGC 716  
QY 77 GCCTCATATCAGCGCGCGTTATCTCATGCGCGGTGACCGGACGAGCAG 122  
DB 717 GTTGCACTTGCGCATCAATCTCAAGTCCGAGAAAGTACAAG 762

RESULT 14  
US-11-097-143-38495  
Sequence 38495, Application US/11097143  
Publication No. US20050208558A1  
GENERAL INFORMATION:



```
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 38495
; LENGTH: 1342
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-38495
```

```
Query Match 17.4%; Score 30.8; DB 13; Length 1342;
Best Local Similarity 55.7%; Pred. No. 1.2;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
Qy 17 ACCGCACTTCTGTGTGTAACCGACGCGGACGCACTCTTATCGGAACGACGC 76
Db 731 ACAACAGCTTCCGGTGAAGTGCCACAAAGGGGTCTTGGCGGCTCCGGTGGC 790
Qy 77 GCCTCATATCAGCGCGGCTTATCTCATGCGCGTGACCGGACAG 122
Db 791 GTTGGCACTTGGCCATACAAATCTCAAGTCCAGAAAGTACAG 836
```

```
RESULT 15
US-11-097-143-38494/c
; Sequence 38494, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
```

```
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 38494
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-38494
```

```
Query Match 17.4%; Score 30.8; DB 13; Length 3448;
Best Local Similarity 55.7%; Pred. No. 1.5;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
Qy 17 ACCGCACTTCTGTGTGTAACCGACGCGGACGCACTCTTATCGGAACGACGC 76
Db 1667 ACAACAGCTTCCGGTGAAGTGCCACAAAGGGGTCTTGGCGGCTCCGGTGGC 1608
Qy 77 GCCTCATATCAGCGCGGCTTATCTCATGCGCGTGACCGGACAG 122
Db 1607 GTTGGCACTTGGCCATACAAATCTCAAGTCCAGAAAGTACAG 1562
```

```
Search completed: June 15, 2006, 18:02:55
Job time : 611.121 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 17:21:03 ; Search time 19.587 Seconds  
(without alignments)  
2127.111 Million cell updates/sec

Title: US-09-896-888a-1\_COPY\_351\_527

Perfect score: 177  
Sequence: 1 gcttctatcgtagacagcgc.....gccgcgaagctctgtttaa 177

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA New:  
1: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
2: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /EMC\_Celerra\_SIDS3/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	30.2	17.1	1137	6	US-10-449-902-2563 Sequence 2563, Ap
C 2	30.2	17.1	1128	6	US-10-449-902-19810 Sequence 19810, A
C 3	30.2	17.1	2019	6	US-10-449-902-22046 Sequence 22046, A
C 4	30.2	17.1	2127	6	US-10-449-902-9697 Sequence 9697, Ap
C 5	30	16.9	1114	6	US-10-953-349-32119 Sequence 32119, A
C 6	30	16.9	3107	6	US-10-517-441-401 Sequence 401, App
C 7	30	16.9	3501	6	US-10-517-441-499 Sequence 499, App
C 8	29.6	16.7	2214	6	US-10-449-902-25649 Sequence 25649, A
C 9	29.6	16.7	2586	6	US-10-449-902-24690 Sequence 24690, A
C 10	29.6	16.7	2618	6	US-10-449-902-16689 Sequence 16689, A
C 11	27.6	15.6	788	6	US-10-449-902-14247 Sequence 14247, A
C 12	27.2	15.4	1687	6	US-10-449-902-25610 Sequence 25610, A
C 13	27.2	15.4	2150	6	US-10-449-902-12192 Sequence 12192, A
C 14	26.8	15.1	1710	7	US-11-217-529-1986 Sequence 1986, Ap
C 15	26.8	15.1	1774	6	US-10-449-902-16716 Sequence 16716, A
C 16	26.8	15.1	1816	7	US-11-293-697-831 Sequence 831, App
C 17	26.8	15.1	2911	7	US-11-293-697-1360 Sequence 1360, App
C 18	26.8	15.1	3020	6	US-10-505-928-320 Sequence 320, App
C 19	26.6	15.0	2791	6	US-10-449-902-17961 Sequence 17961, A
C 20	26.4	14.9	974	6	US-10-449-902-1273 Sequence 1273, Ap
C 21	26.4	14.9	1092	7	US-11-217-529-77591 Sequence 77591, A
C 22	26.4	14.9	1848	6	US-10-449-902-376 Sequence 376, App
C 23	26.4	14.9	1848	6	US-10-449-902-8998 Sequence 8998, App
C 24	26.4	14.9	3169	6	US-10-449-902-12991 Sequence 12991, A
C 25	26	14.7	2768	6	US-10-449-902-19286 Sequence 19286, A

C 26	25.8	14.6	1439	6	US-10-449-902-21106 Sequence 21106, A
C 27	25.8	14.6	4119	7	US-11-217-529-77046 Sequence 77046, A
C 28	25.6	14.5	1775	6	US-10-449-902-81373 Sequence 81373, Ap
C 29	25.6	14.5	2154	6	US-10-449-902-14093 Sequence 14093, A
C 30	25.6	14.5	2627	6	US-10-449-902-21824 Sequence 21824, A
C 31	25.6	14.5	2691	6	US-10-449-902-25676 Sequence 25676, A
C 32	25.6	14.5	3558	7	US-11-121-154-173 Sequence 173, Appl
C 33	25.4	14.4	1403	7	US-11-313-836-43 Sequence 43, Appl
C 34	25.4	14.4	1476	7	US-11-217-529-681 Sequence 681, Appl
C 35	25.4	14.4	1688	6	US-10-953-349-31139 Sequence 31139, A
C 36	25.4	14.4	3061	6	US-10-449-902-12832 Sequence 12832, A
C 37	25.4	14.4	3088	6	US-10-449-902-12783 Sequence 12783, A
C 38	25.2	14.2	1644	6	US-10-449-902-27203 Sequence 27203, A
C 39	25.2	14.2	1773	6	US-10-449-902-476 Sequence 476, App
C 40	25.2	14.2	1782	6	US-10-449-902-22687 Sequence 22687, A
C 41	25.2	14.2	2207	6	US-10-449-902-17975 Sequence 17975, A
C 42	25	14.1	1924	6	US-10-449-902-22070 Sequence 22070, A
C 43	25	14.1	2619	6	US-10-449-902-18809 Sequence 18809, A
C 44	25	14.1	3310	6	US-10-449-902-12737 Sequence 12737, A
C 45	24.8	14.0	937	6	US-10-449-902-1893 Sequence 1893, Ap

#### ALIGNMENTS

RESULT 1  
US-10-449-902-2563/C  
Sequence 2563, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449, 902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 2563  
LENGTH: 1127  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK104298  
DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-2563

Query Match 17.1%; Score 30.2; DB 6; Length 1127;  
Best Local Similarity 53.9%; Pred. No. 0.15; Mismatches 53; Indels 0; Gaps 0;  
Matches 62; Conservative 0; Mismatch 53; Indel 0; Gap 0;

Qy 16 GAGCCAGCTTCTGTGTGCTTAACCGACGCGGACGCACTCTTATCGAAGAGAGG 75  
Db 199 GCGCGCGGTGTCGCGCATGCGCGCGCGCGCGCGCGCATGCGCGCGCGCGCGCG 140

Qy 76 CCGCTTCATATCAGCGCGCGCGGTATTCATGCGCGGTGACCGGACGAGAGAGCGCC 130  
Db 139 CCGCGCGCTCTCTCCGACGCGATCGACGCGCGCGCGCGCGCGGTGACGAGAGCGCC 85

RESULT 2  
US-10-449-902-19810/C  
Sequence 19810, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.

APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449, 902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 19810  
LENGTH: 1128  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK070153  
DATABASE ENTRY DATE: 2001-12-06  
US-10-449-902-19810

Query Match 17.1%; Score 30.2; DB 6; Length 1128;  
Best Local Similarity 53.9%; Pred. No. 0.15; Indels 0; Gaps 0;  
Matches 62; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 16 GACGCGAGCTTCTGCTGTGCTAACGCGAGCGGCACTCTTATCGGAACAGAGCG 75  
DB 200 GCGGCGCGCTCTGCTGCGGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 141  
QY 76 GCGCTCATATCAGCGCGCGCTTATCTCATGCGCGTACCGGACAGAGCGCGCC 130  
DB 140 CCCCCGCTCTCTCCGACGCGCATCCGAGCGCGGCGCTCGGTACGATGAGGCC 86

RESULT 3  
US-10-449-902-22046  
Sequence 22046, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449, 902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 22046  
LENGTH: 2019  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK072256  
DATABASE ENTRY DATE: 2001-12-06  
US-10-449-902-22046

Query Match 17.1%; Score 30.2; DB 6; Length 2019;  
Best Local Similarity 55.1%; Pred. No. 0.18; Indels 48; Gaps 0;  
Matches 59; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 38 AACGCGAGCGGAGCGCACTCTTATCGGAACAGAGCGCGCTTATCGAGCGCGCT 97  
DB 1562 AACGTCATCAGAGCGCGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 1621  
QY 98 TATCTATGCGCGGACCGGACAGAGCGCGCGCTCCGCTTATCGC 144  
DB 1622 GACCTATCGCGCGCTGCTCTGTCAGAGAGCCCAAGAGCGGATCGC 1668

RESULT 4  
US-10-449-902-9697  
Sequence 9697, Application US/10449902  
Publication No. US20060123505A1  
GENERAL INFORMATION:  
APPLICANT: National Institute of Agricultural Sciences.  
APPLICANT: Bio-oriented Technology Research Advancement Institution.  
APPLICANT: The Institute of Physical and Chemical Research.  
APPLICANT: Foundation for Advancement of International Science.  
TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF  
FILE REFERENCE: MOA-A0205Y1-US  
CURRENT APPLICATION NUMBER: US/10/449, 902  
CURRENT FILING DATE: 2003-05-29  
PRIOR APPLICATION NUMBER: JP 2002-203269  
PRIOR FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: JP 2002-383870  
PRIOR FILING DATE: 2002-12-11  
NUMBER OF SEQ ID NOS: 56791  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 9697  
LENGTH: 2127  
TYPE: DNA  
ORGANISM: Oryza sativa  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AK107045  
DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-9697

Query Match 17.1%; Score 30.2; DB 6; Length 2127;  
Best Local Similarity 55.1%; Pred. No. 0.18; Indels 48; Gaps 0;  
Matches 59; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 38 AACGCGAGCGGAGCGCACTCTTATCGGAACAGAGCGCGCTTATCGAGCGCGCT 97  
DB 1543 AACGTCATCAGAGCGCGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 1602  
QY 98 TATCTATGCGCGGACCGGACAGAGCGCGCGCTCCGCTTATCGC 144  
DB 1603 GACCTATCGCGCGCTGCTCTGTCAGAGAGCCCAAGAGCGGATCGC 1649

RESULT 5  
US-10-953-349-32119/C  
Sequence 32119, Application US/10953349  
Publication No. US20060107345A1  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai et al.  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
FILE REFERENCE: 2750-1579PUS2  
CURRENT APPLICATION NUMBER: US/10/953, 349  
CURRENT FILING DATE: 2004-09-30  
NUMBER OF SEQ ID NOS: 40252  
SOFTWARE: Patentin version 3.3  
SEQ ID NO 32119  
LENGTH: 1114  
TYPE: DNA  
ORGANISM: Trilicium aestivum  
US-10-953-349-32119

Query Match 16.9%; Score 30; DB 6; Length 1114;  
Best Local Similarity 53.4%; Pred. No. 0.17; Indels 55; Gaps 0;  
Matches 63; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 29 TGTGTGCTAACGCGAGCGCACTCTTATCGGAACAGAGCGCGCTTATCGA 88  
DB 955 TGTGTGCGATGACAG 896  
QY 89 GCGCGCGCTTATCTCATGCGCGTACCGGACAGAGCGCGCGCTCCGCTTATCGCGC 146  
DB 895 GCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838

```
RESULT 6
US-10-517-441-401/c
; Sequence 401, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOCKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almut
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 401
; LENGTH: 3107
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-401

Query Match      16.9%; Score 30; DB 6; Length 3107;
Best Local Similarity 50.7%; Pred. No. 0.23; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 70;

QY 6 ATGTGACAGAGAGCCAGCTTCTGTGTCCTAACCGAGCGAGCACTCTATCG 65
DB 2079 AACGGAACACACCCGCTCTCTCGCTCCAAACCGCGCGAATTAATCAACCGCC 2020
QY 66 GAACAGAGCGCGCTCTCATATTCAGCGCGGTTATCTCATGCGCGTACCGAGG 125
DB 2019 GAACCGCGAGCTCTCTCTTTTAAACAATCTTAAGCGCGCGAACAAGACCGCC 1960
QY 126 CGCCCGTCCGCTTATCGCGCC 147
DB 1959 CACTTCTCCCATATCGCGCC 1938

RESULT 7
US-10-517-441-499/c
; Sequence 499, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOCKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
```

```
APPLICANT: SCHMITT, Manfred
APPLICANT: LOOK, Maxime P.
APPLICANT: MARX, Almut
APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 499
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-499

Query Match      16.9%; Score 30; DB 6; Length 3501;
Best Local Similarity 50.7%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 70;

QY 6 ATGTGACAGAGAGCCAGCTTCTGTGTCCTAACCGAGCGAGCACTCTATCG 65
DB 1585 AACGGAACACACCCGCTCTCTCTCGCTCCAAACCGCGCGAATTAATCAACCGCC 1526
QY 66 GAACAGAGCGCGCTCTCATATTCAGCGCGGTTATCTCATGCGCGTACCGAGG 125
DB 1525 GAACCGCGAGCTCTCTCTTTTAAACAATCTTAAGCGCGCGAACAAGACCGCC 1466
QY 126 CGCCCGTCCGCTTATCGCGCC 147
DB 1465 CACTTCTCCCATATCGCGCC 1444

RESULT 8
US-10-449-902-25649
; Sequence 25649, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25649
; LENGTH: 2214
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK101091
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-25649

Query Match      16.7%; Score 29.6; DB 6; Length 2214;
Best Local Similarity 52.4%; Pred. No. 0.29;
```



```

Qy      28  CTGTTGTCCTAACCGCAGCCGGAAGCAACTCTCTTATCGGAACAGAGCGGCGCTTCATATC  87
Db      694  CGGGCTCCCACTCGGCGCGCCACCCGGAAGTCCATGCGGATCGGAACCCGAGCCAGTAGC  635

Qy      88  AGCCGCGCGCTTATCTCATGCGCGTAGCCGGAACAGAGCGCGCCG  131
Db      634  TGTGAACGCGGCGGACCTGCGCGTGGGGGGAATCGAGGTTCCCG  591

RESULT 14
US-11-217-529-1986
; Sequence 1986, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1986
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1986

Query Match      15.1%; Score 26.8; DB 7; Length 1710;
Best Local Similarity 49.3%; Pred. No. 2.6;
Matches 70; Conservative 0; Mismatches 72; Indels 0; Gaps 0

Qy      7  TCGTAGCAGAGCGCAGCTTCCTGTGTGCTPAACCGCAGCGGAGCAACTCTTATCGG  66
Db      210  TCTTCACAAATGTCGTGCTATATAGTGTGCCAAGGGGTCAAAAAGTAGTATTGGCCA  269

Qy      67  AACAGAGCGCGCTTCATATAGCGCGCGTTATCTATGCGCGTAGACCGACAGAGCG  126
Db      270  AAAAGGTCCTTATCTACACAGCGCGCTTTCACATATGAGAACTTATGAGCAAAATG  329

Qy      127  GCCGTCGCCGCTTATCGGCGCT  148
Db      330  CACCGCGGATTAATCTTGACT  351

RESULT 15
US-10-449-902-16716
; Sequence 16716, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16716
; LENGTH: 1774
; TYPE: DNA

```

```

; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK067192
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-16716

```

```

Query Match      15.1%; Score 26.8; DB 6; Length 1774;
Best Local Similarity 64.5%; Pred. No. 2.6;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY      76 CGCCTCCATATCAAGCGCGCTTATCTCATGCGCGTGACCGGACACGAGGCGCCCGTCCC 135
Db      70 CGGCTTCGCTCAAGCTTCGCTCTCTCGCGCGCGCGCGCGGAGCGGAGGACACGAGTGC 129

```

```

QY      136 GC 137
Db      130 GC 131

```

```

Search completed: June 15, 2006, 18:04:17
Job time : 20.587 secs

```